NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-59, Vol. III

Supporting Research and Advanced Development

For the Period August 1 to September 30, 1969

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

October 31, 1969

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-59, Vol. III

Supporting Research and Advanced Development

For the Period August 1 to September 30, 1969

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

October 31, 1969

SPACE PROGRAMS SUMMARY 37-59, VOL. III

Prepared Under Contract No. NAS 7-100
National Aeronautics and Space Administration

Preface

The Space Programs Summary is a multivolume, bimonthly publication that presents a review of technical information resulting from current engineering and scientific work performed, or managed, by the Jet Propulsion Laboratory for the National Aeronautics and Space Administration. The Space Programs Summary is currently composed of four volumes.

Vol. I. *Flight Projects* (Unclassified)

Vol. II. *The Deep Space Network* (Unclassified)

Vol. III. *Supporting Research and Advanced Development* (Unclassified)

Vol. IV. *Flight Projects and Supporting Research and Advanced Development* (Confidential)

Contents

SPACE SCIENCES DIVISION

TELECOMMUNICATIONS DIVISION

Contents (contd)

GUIDANCE AND CONTROL DIVISION

Contents (contd)

Contents (contd)

Contents (contd)

MISSION ANALYSIS DIVISION

Contents (contd)

I. Lunar and Planetary Instruments

SPACE SCIENCES DIVISION

A. Increased Sensitivity to Creatinine With Ethanol, *H. Y. Tom, L. Taylor, and D. Alexander*

1. Introduction

The Jet Propulsion Laboratory is responsible for the development of a urine analyzer for the quantitative determination of creatinine, creatine, and calcium as part of the Biosatellite Program (Ref. 1). The mechanical and electrical details of this instrument are described in Ref. 2 and the chemistry is to be published at a later date.[1] This article demonstrates that the Jaffe reaction (Ref. 3) applied by O. Z. Folin (Ref. 4), and modified at Jet Propulsion Laboratory,[1] can be made more sensitive to creatinine by the employment of ethanol as a solvent for picric acid crystals.

2. Method

Picric acid crystals were purified using the method developed by J. Hardy.[2] Briefly, this method consists of boiling a solution of 220 g of crystals and 440 ml of H_2O in an Erlemmeyer flask. Enough water is added to dissolve the crystals and the solution is then filtered through a preheated Buchner funnel into a vacuum filtration flask. The filtered solution is again heated to drive off the water, and the crystals are then dissolved, reheated to boiling, and refiltered. The filtrate is allowed to cool, whereupon a supersaturated solution, used as a stock solution, is obtained. A similar procedure was employed using ethanol for the 70 vol % ethanol picric acid solution.

The absorbance from an aliquot of solution was measured with either a Cary 14 or a Beckman spectrophotometer, Model DB, at 490 nm wavelength. The aliquots came from two different formulations, the difference between them being that H_2SO_4 was used in one of them for the purpose of converting creatine to creatinine (Footnote 1). The solution prepared without H_2SO_4 was made by using 0.100 ml of creatinine sample, 2.00 ml of 0.03 N picric acid using H_2O or 70 vol % ethanol as its solvent, 1.00 ml of either 2.00 N or 2.25 N NaOH, and then to 10 ml total volume with H_2O. The reagent blank contained the same reagents as the sample except with 1.000 ml H_2O. The absorbance of the unknown compared to the reagent blank was determined with a Cary 14.

[1]Personal communication from J. Rho of the JPL Bioscience Section.

[2]Personal communication from J. Hardy of the JPL Bioscience Section.

The other formulation is prepared by taking 0.300 ml from a solution of 0.300 ml sample and 1.62 ml of 4 N H_2SO_4 and adding it to 0.865 ml of NaOH and 0.755 ml of 0.030 picric acid. The aliquot from this solution was also read at 490 nm using a reagent blank in the reference beam. A Beckman spectrophotometer was used for these determinations.

3. Results

The results from the samples not using the H_2SO_4 formulation are displayed in Fig. 1. In Fig. 1, curves A and B have been combined because there is no apparent difference between them within the error of this experiment. The curves shown reflect the different chemistry of the basicity of the solutions and whether they use ethanol picric acid or aqueous picric acid. That is, the composite A and B curve includes aqueous picric acid with 2.00 N NaOH for the former and 2.25 N NaOH for the latter; curves C and D contain 70 vol % ethanol solvent for picric acid with 2.25 N NaOH for curve C and 2.00 N NaOH for curve D.

One immediately obvious conclusion that may be drawn from Fig. 1 is that ethanol picric acid enhances

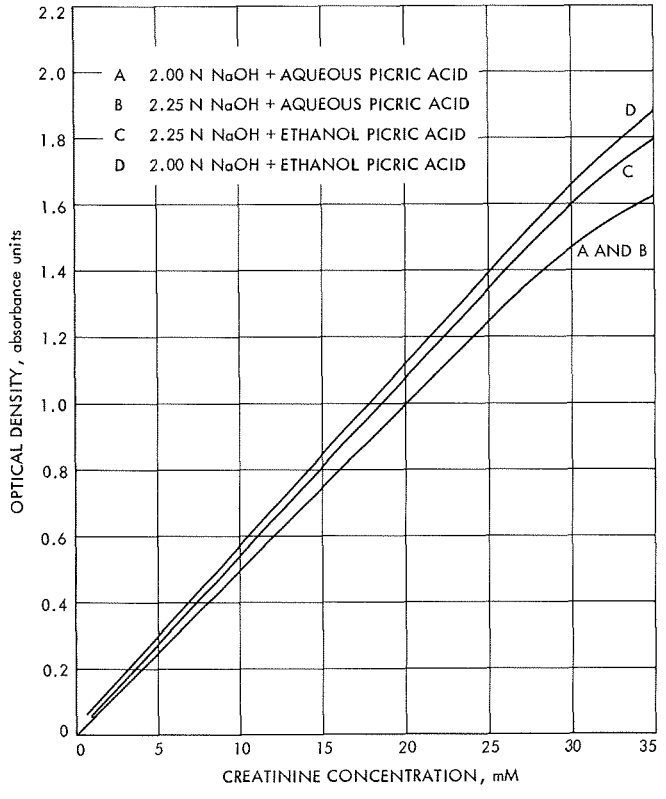


Fig. 1. Absorbance of creatinine without H_2SO_4

the Jaffe reaction. The conclusion is supported by the greater slope of curves C and D as compared to composite curve A and B. The only difference between the two sets of curves is the presence of ethanol for the former and the absence of ethanol for the latter.

Any major influence from basicity is not evident when one employs aqueous picric acid in the reaction; it is only at 35 mM creatinine concentrations that there is some 0.05 absorbance units of difference. With ethanol picric acid, however, the less alkaline solution (Fig. 1, curve D) has a greater slope than a more concentrated solution of NaOH (Fig. 1, curve C). Thus, a less basic condition has the tendency to enhance the Jaffe reaction for creatinine.

In Fig. 2, the results from those solutions that incorporate H_2SO_4 show the tendency to make the Jaffe method more suitable for the quantitative determination of creatinine by comparing curves A and B where the only chemical difference is that curve B is from a solution that is closer to neutrality than curve A. Moreover, the more neutral solution also tends to increase the range of the method since the curvature changes at higher creatinine concentrations, and, in this case, 13.5 mM of creatinine compare to 10 mM having a 2 N NaOH for the former and 2.25 N NaOH for the latter.

Although alkaline picrate, with different basicity was used to generate curves B and C of Fig. 2, curve C has a much wider range and is linear throughout the range that was investigated. Thus, even though a more alkaline picrate solution was used by the employment of ethanol as a solvent for picric acid, the Jaffe method may be greatly improved by the application of alcohol as a solvent for picric acid.

4. Conclusions

It has been conclusively demonstrated that if one wishes to use the Jaffe reaction for the determination of creatinine, one should employ ethanol as a solvent for picric acid. However, under conditions of low temperatures, this method becomes unsatisfactory because the picric acid crystals precipitate in ethanol, but not in water, at 16°C. Thus, one may use an ethanol solvent for laboratory applications but its employment becomes unsatisfactory at reduced temperatures.

References

1. NASA 7-100, Task Order RD-64, National Aeronautics and Space Administration, Washington.

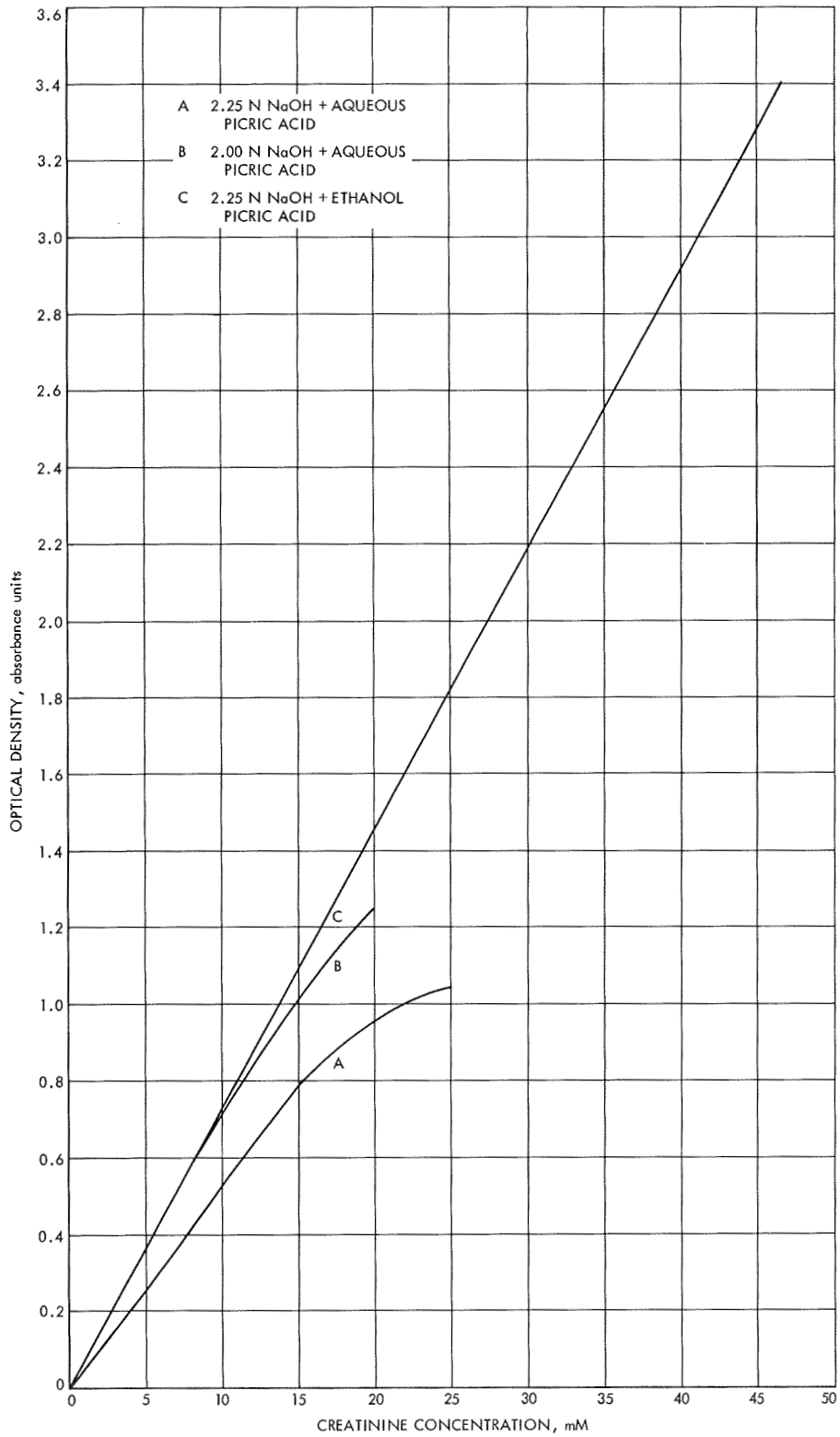


Fig. 2. Absorbance of creatinine with H_2SO_4

2. Stuart, J. L., paper presented at the third annual meeting of the Association for the Advancement of Medical Instruments, Houston, Texas, July 16, 1969.
3. Jaffe, M. Z., *Physiol. Chem.*, Vol. 10, p. 391, 1886.
4. Folin, O. Z., *Physiol. Chem.*, Vol. 41, p. 223, 1904.

B. Biosatellite Urinalysis Instrument: The Data Amplifiers, *J. R. Clark*

1. Introduction

The Biosatellite urinalysis instrument required two data amplifiers: one for calcium measurement and one for the creatinine measurement. They are both used to process the signal resulting from a change in resistance of photoconductive cells. The calcium data amplifier processes a signal proportional to the amount of fluorescence resulting from the degree of calcium concentration in the fluid being analyzed. The creatinine amplifier produces a response proportional to the optical density of the fluid being analyzed. The spacecraft interface required that the output voltage be 0–5 V. It was arbitrarily established that the voltage increase for increasing concentration.

2. Creatinine Data Amplifier

a. Description. The creatinine data amplifier views light with a center wavelength of 500 nm from the optical assembly of the creatinine analyzer. This light falls on a photoconductive cell having a spectral response as shown in Fig. 3. The photocell is the feedback element around an operational amplifier that causes the output voltage of the data amplifier to vary with the concentrations of creatinine present in the test cell. Figure 4 traces the signal flow from concentration to output voltage. The details of the optical system have been omitted.

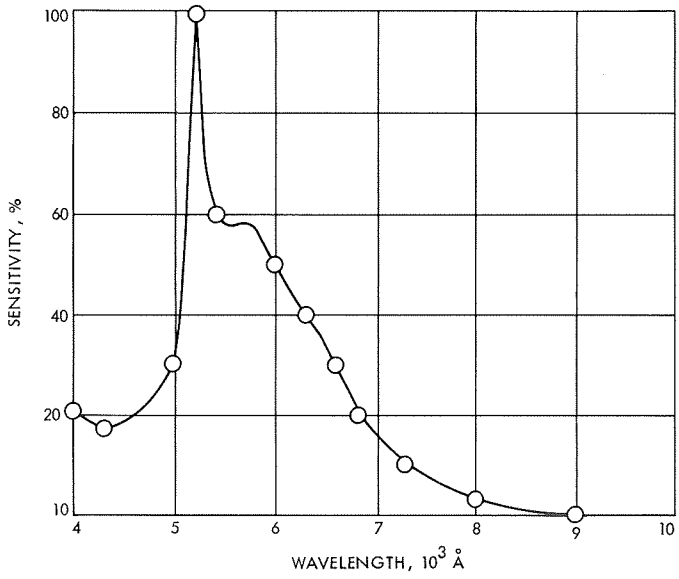


Fig. 3. CdS photocell (Clairex 903N) spectral response

Working on an incremental basis, an increase in concentration of creatinine, due to the chemical reaction in the test cell, causes the transmitted light intensity to decrease. A decreased light intensity on the photocell causes its resistance to increase. This increased cell resistance causes an increase in output voltage. The output voltage increases because the amplifier is biased to operate in the second quadrant. This sequence is shown below.

$$Cr_{\text{conc}} \uparrow \qquad I \downarrow \qquad R_T \uparrow \qquad E_0 \uparrow$$

where

Cr_{conc} = creatinine concentration

I = light intensity

R_T = test photocell resistance

E_0 = amplifier output voltage

A reference photocell viewing a reference cell is used as the input resistor in Fig. 4. Optically, this cell is a blank in that only reagents are added to it. A common light source is used, which tends to cancel this common-mode effect.

b. Circuit configuration. Figure 5 shows the detailed circuit as it presently exists in Unit 005. Some of the

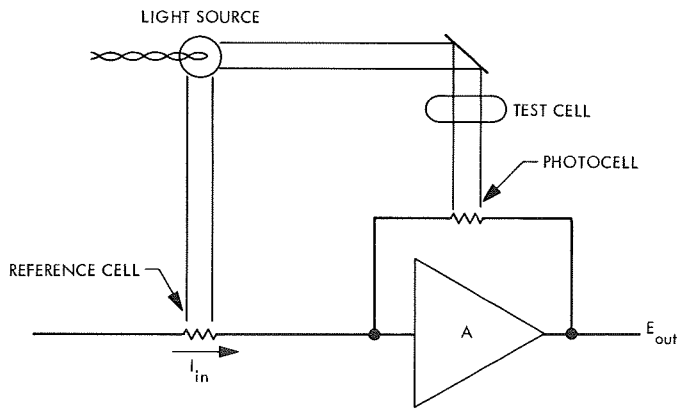


Fig. 4. Creatinine signal diagram

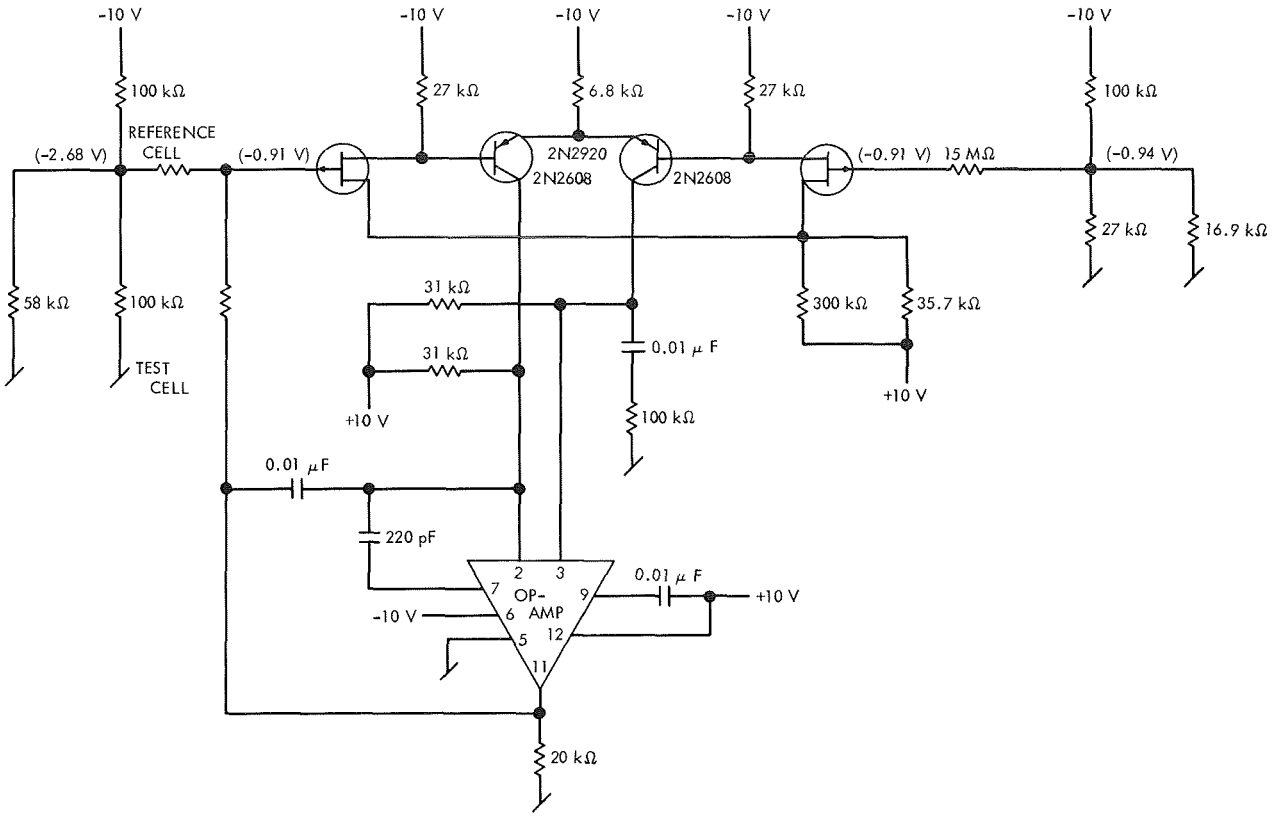


Fig. 5. Creatinine data amplifier circuit diagram

resistances are tailored so that maximum temperature stability can be achieved.[3]

Figure 6 shows the assembled circuit board. The amplifier occupies the upper half of the circuit board. The large 50-MΩ resistors have been changed to the values shown on Fig. 5. The integrated circuit amplifier is located on the right side of the circuit board. The trim pot shown in the upper right-hand corner has been deleted.

c. Data amplifier biasing considerations. The input field-effect transistors used to increase input resistance are of primary concern since they are the major contributors to output voltage drift due to local and ambient temperature changes. This is because this input stage is not preceded by any voltage gain, i.e., any input changes due to temperature fluxuations will appear at the output just as though there had been a change in the reference voltages. For these reasons, matched field-effect transistors operating in a differential mode were used as the input elements to minimize the initial offset of the gate-source voltage V_{GS} ($V_{GS1} - V_{GS2}$). Changes in the transistors drain current are minimized by biasing

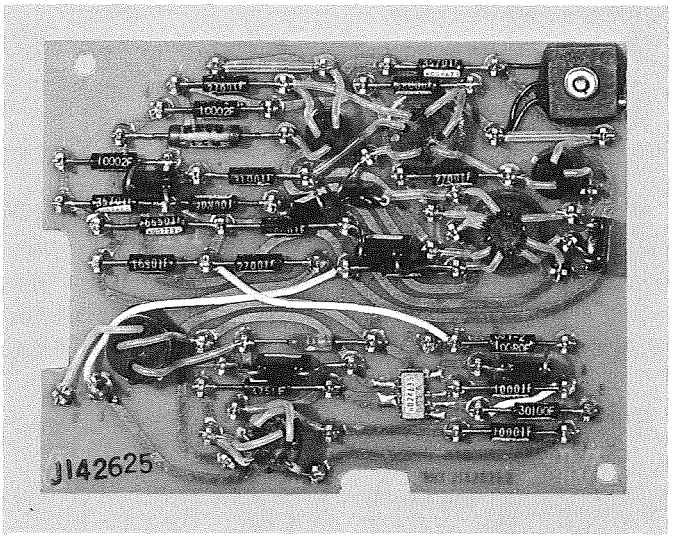


Fig. 6. Creatinine data amplifier flight configuration

[3]Fujimoto, H., *Acceptance Test Procedure CSP 142625 Biosatellite Component Selection Procedure for Creatinine Data Amplifier* (JPL internal document).

each field-effect transistor at its zero temperature coefficient location. This can be done since the V_{GS} and transconductance have opposite temperature coefficients. Since the transistors are matched, this bias point is about the same for both field-effect transistors. For the devices used, the zero temperature coefficient point occurs when the V_{GS} is +1.5 V and the drain current is about 150 μA. The V_{GS} were matched to less than 10 mV. The chosen bias point for the field-effect transistors sets the quiescent V_{GS} and drain voltage. The dual NPN transistor is used to level shift the drain voltage to 0 V, which is the optimum bias voltage for driving the Amelco operational amplifier. Table 1 lists the operating characteristics of the integrated-circuit operational amplifier used.

Table 1. Integrated circuit op-amp operating characteristics (Type: A13-251-Amelco)

Open loop dc gain (no load)	20,000 V (typical)
Open loop bandwidth	10 MHz
Input offset voltage	5 mV
Input current	500 nA
Input offset voltage drift	25 V/°C
Input current drift (−55 to +125°C)	5 nA/°C
Input inpedance	1 MΩ
Output inpedance (open loop)	750 Ω
Common mode input range	± 2.0 V
Dynamic output range (no load)	18 V (p–p)
Dynamic output range (1kΩ load)	10 V (p–p)
Common mode rejection	−80 dB

d. DC analysis–transfer relationships. Figure 7 shows the equivalent circuit that was used to derive expressions for reference-voltage-to-output-voltage and test-cell-resistance-to-output-voltage. For these deviations, the gain A of the operational amplifier is assumed infinite.

Deviation of voltage transfer functions

$$E_0 = E_2\left(\frac{1}{\frac{R_R}{R_R+R_T}-\frac{1}{A}}\right) - E_1\left(\frac{1}{\frac{R_R}{R_T}-\frac{R_R+R_T}{AR_T}}\right) \qquad (1)$$

where R_R is the reference cell resistance. And, when $A \gg 1$,

$$E_0 = E_2\left(\frac{R_T+R_R}{R_R}\right) - E_1\left(\frac{R_T}{R_R}\right) \qquad (2)$$

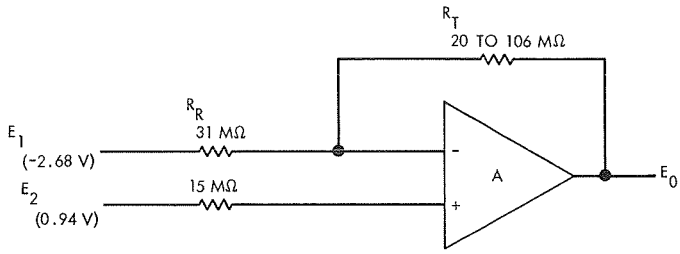


Fig. 7. Transfer function derivation diagram

Equation (1) places an arbitrary upper limit on R_T. The dc gain of the amplifier should be high enough so that the error terms $1/A$ and

$$\frac{1}{A}\frac{R_R+R_T}{R_T}$$

have negligible effect on the output voltage.

Intercept derivations of Eq. (2). By knowing the intercepts of the function, E_0 as a function of E_1 can be easily plotted. These intercepts are obtained from Eq. (2).

When

$$E_1 = 0, \qquad E_0 = E_2\frac{R_T+R_R}{R_R} \qquad (3)$$

When

$$E_0 = 0, \qquad E_1 = \frac{R_R}{R_T}\left(\frac{R_T+R_R}{R_R}\right)E_2 \qquad (4)$$

Q-point derivation. Another useful point on the voltage-transfer curve is the Q point, which is the point on the curve where changes in test cell resistance will cause no change in output voltage. It is defined by the following derivation. From Eq. (2),

$$\frac{\partial E_0}{\partial R_T} = 0 = \frac{E_2}{R_R} - \frac{E_1}{R_R} \qquad (5)$$

or it occurs when $E_2 = E_1$. Then, when $E_2 = E_1$, E_0 is found from Eq. (2) as $E_0 = E_1 = E_2$. This Q point is important because it does not change during operation of the amplifier while the intercepts will, because they are a function of the test cell resistance.

Figure 8 is a plot of the transfer characteristic showing slope, intercepts, and Q point. It is plotted in the first

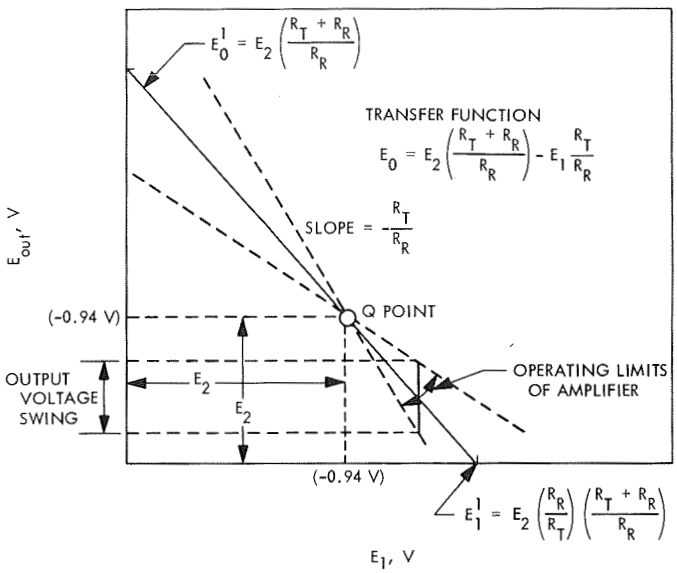


Fig. 8. Voltage transfer characteristic

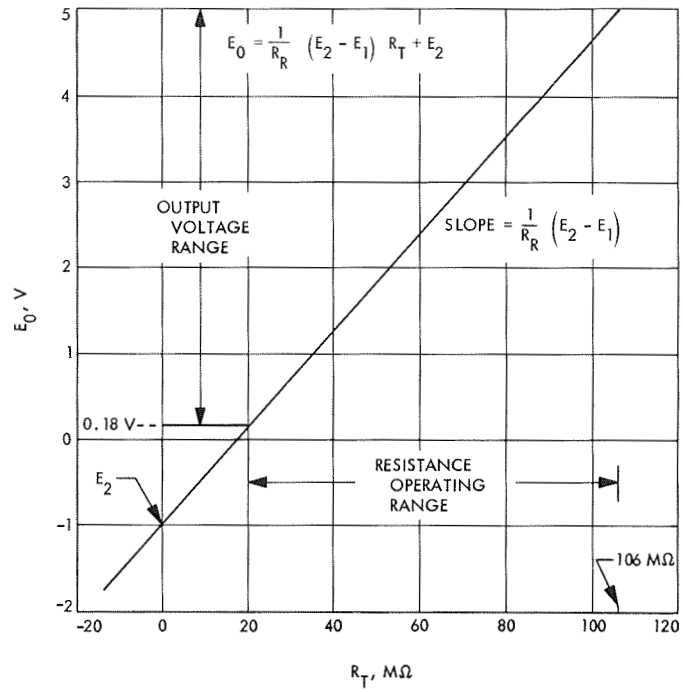


Fig. 9. Creatinine data amplifier output voltage versus test cell resistance

quadrant, although, in actuality, the amplifier is operating in the second quadrant because E_1 and E_2 are negative.

Deviation of output voltage versus test cell resistance. From Eq. (2),

$$\frac{\partial E_0}{\partial R_T} = \frac{1}{R_R}(E_2 - E_1)$$

and, when $R_T = 0$, $E_0 = E_2$ (from Eq. 5). Then, using the slope–intercept form to define the characteristic,

$$E_0 = \frac{1}{R_R}(E_2 - E_1)R_T + E_2 \qquad (6)$$

This function is plotted in Fig. 9 using the following values for the variable:

$$R_R = 31\ \text{M}\Omega$$
$$E_2 = 0.94\ \text{V}$$
$$E_1 = 2.68\ \text{V}$$
$$R_T = 20\text{–}106\ \text{M}\Omega$$

3. Calcium Data Amplifier

a. Description. This amplifier views light from the calcium fluorometer. The light falls on a photoconductive cell having a spectral response as shown in Fig. 3. This photocell is the feedback element around an operational amplifier. Its resistance changes from 133 to 270 MΩ.

Figure 10 will be useful in following concentration changes in the test cell to output voltage changes of the amplifier.

Working on an incremental basis, an increase in the concentration of calcium increases the fluorescence in the reaction and, therefore, the transmitted light intensity increases. This causes the resistance of the test photocell to decrease, which causes the output voltage of the

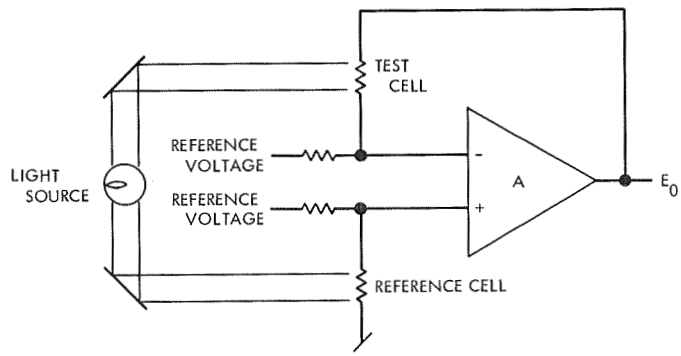


Fig. 10. Calcium signal diagram

amplifier to increase. This voltage increases because the amplifier is biased to operate in the first quadrant rather than the fourth and second. This results in an output voltage that is proportional to the calcium concentration. These events are shown below.

$$Ca^{++}_{\text{conc}} \uparrow \qquad I \uparrow \qquad R_T \downarrow \qquad E_0 \uparrow$$

where Ca^{++}_{conc} is the calcium concentration.

The complete amplifier schematic is shown in Fig. 11. Since the source impedance is fairly high, two field-effect transistors operating as a differential input stage were used. Since the fluorometer contained a reference cell, it was connected as shown in Figs. 10 and 11 to reduce drift at the output caused by changing source intensity. The two field-effect transistors are biased such that their drains are close to ground to provide optimum coupling to the integrated-circuit operational amplifier. Other bias considerations of the field-effect transistors are contained in *Subsection 2-c*. Figure 12 shows the flight configuration of the calcium data amplifier. The integrated-circuit operational amplifier operating characteristics are listed in Table 1.

b. Voltage transfer function. The transfer function can be derived from the simplified circuit of Fig. 13.

$$E_0 = E_2\left[\frac{1}{\left(\frac{R_2}{R_1+R_2}\right)\left(\frac{R_3+R_4}{R_4}\right) - \frac{R_3+R_4}{AR_4}}\right] - E_1\left(\frac{1}{\frac{R_2}{R_1} - \frac{R_1+R_2}{AR_1}}\right) \tag{7}$$

And, when $A \gg 1$,

$$E_0 = E_2\left[\left(\frac{R_1+R_2}{R_2}\right)\left(\frac{R_4}{R_4+R_3}\right)\right] - E_1\left(\frac{R_1}{R_2}\right) \tag{8}$$

As stated earlier, the amplifier was designed to operate in the first quadrant. Significant points of the transfer functions (axis intercepts and Q point) are shown in Fig. 14. The Q point is that value of input voltage (E_1, E_2) such that the test cell resistance can be varied without changing the output voltage E_0.

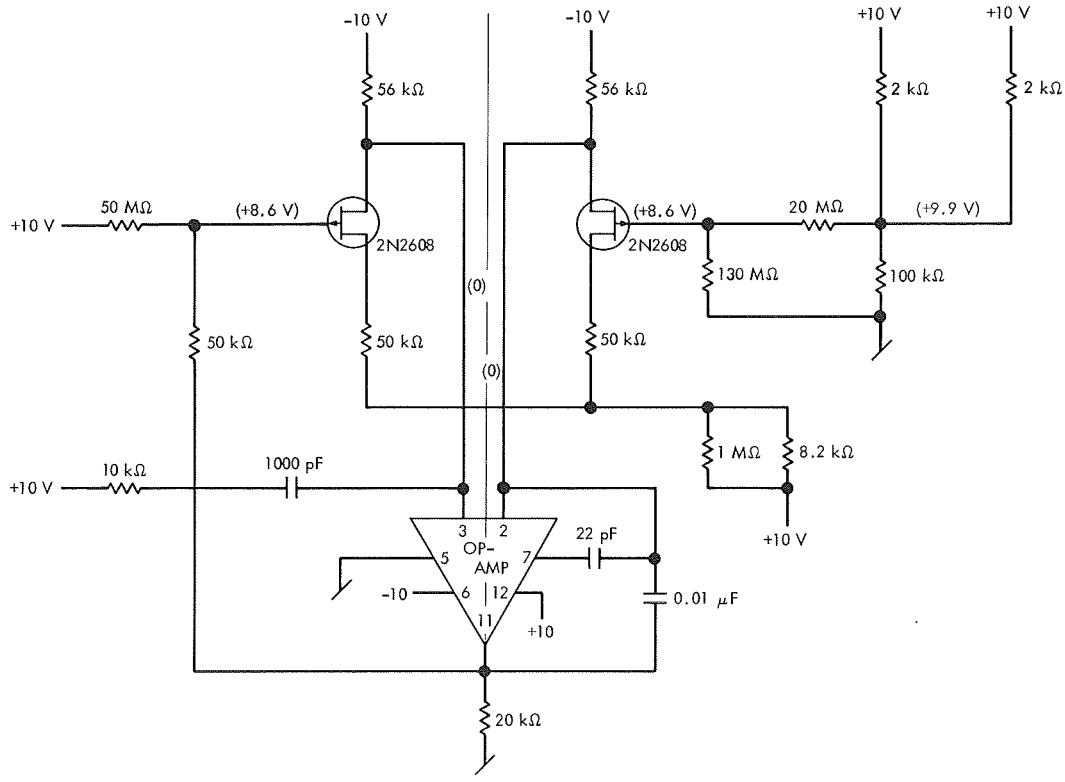


Fig. 11. Calcium data amplifier

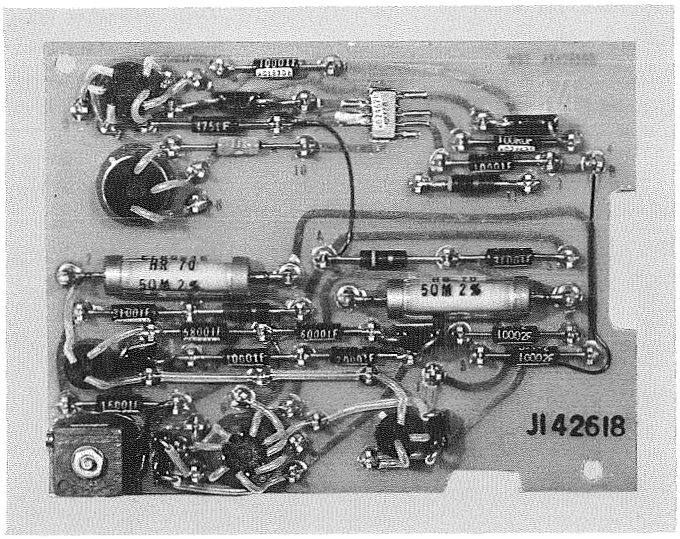


Fig. 12. Calcium data amplifier flight configuration

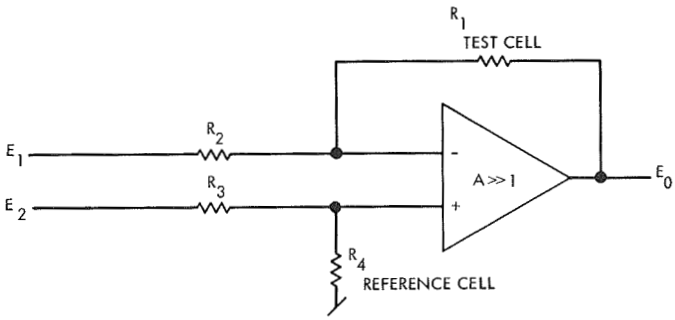


Fig. 13. Calcium data amplifier model

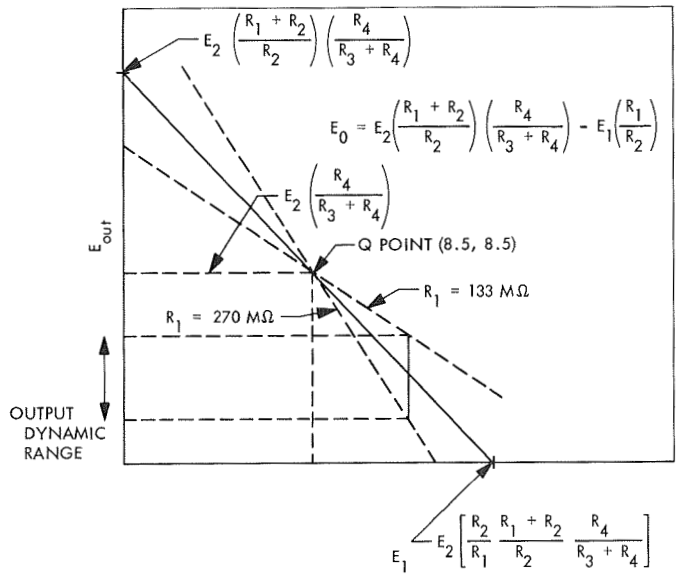


Fig. 14. Transfer characteristic

The test cell resistance transfer curve from the transfer function (Eq. 8) derives E_0 as a function of test cell resistance. This is given below in the slope–intercept form

$$E_0 = \frac{1}{R_2}\left(E_2\frac{R_4}{R_3+R_4} - E_1\right)R_1 + E_2\left(\frac{R_4}{R_3+R_4}\right) \quad (9)$$

With the present design, the voltage and resistance are as follows:

$$E_1 = 10\ \text{V}$$
$$E_2 = 10\ \text{V}$$
$$R_2 = 50\ \text{M}\Omega$$
$$R_3 = 20\ \text{M}\Omega$$
$$R_4 = 130\ \text{M}\Omega$$

Substituting these values into Eq. (9) gives

$$E_0 = R_1\left(-0.057\frac{\text{volts}}{\text{M}\Omega}\right) + 8.66$$

where R_1 is the test cell resistance (Fig. 15).

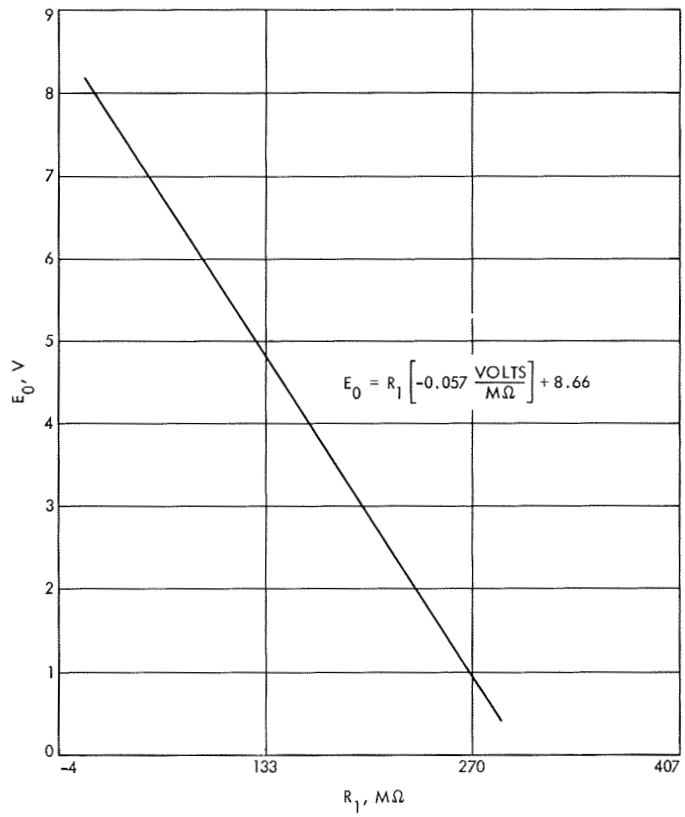


Fig. 15. Calcium data amplifier output voltage versus test cell resistance

Note that, per initial requirements, the slope is negative. That is, as the concentration of calcium increases, the test cell resistance decreases causing the output voltage to increase.

4. Conclusions

Using presently available integrated circuits, the amplifiers could be updated to eliminate the input field-effect and level-shifting transistors required by this design. Operational amplifiers are available from Amelco that include the field-effect transistors as the input section, which would reduce the part count and the space required to mount the amplifier circuit board.

If feasible, it would be desirable to incorporate the whole integrated-circuit data amplifier into the case of the photocell, thereby reducing noise pickup on the input high-impedance leads.

II. Space Instruments

SPACE SCIENCES DIVISION

A. Cryogenically Cooled Preamplifier, *L. L. Lewyn*

1. Introduction

High-resolution solid-state detectors are replacing NaI (Tl) scintillators in many gamma-ray spectroscopy applications. The resolution of a 30-cm^3 coaxial Ge (Li) detector is typically 3 keV. The resolution of a good 3×3 NaI (Tl) scintillator photomultiplier assembly at the 0.662 Mev CS137 energy is approximately 46 keV. Charge sensitive preamplifiers based on bipolar transistor first-stage circuitry are adequate for scintillation spectroscopy applications.

Field effect transistor (FET) first-stage circuitry has been employed in uncooled commercial preamplifiers to reduce preamplifier noise to a level of approximately 2.5 keV. This noise level applies to operation with 30-cm^3 solid-state detectors which have a capacity of approximately 30 pF. At a noise level of 2.5 keV, the preamplifier contribution to system noise is approximately the same as the detector contribution. In order to further reduce preamplifier noise, cryogenic cooling of the first stage must be employed.

2. System Configuration

The preamplifier first stage is electrically and mechanically isolated from the remainder of the circuit. A sketch of the first-stage cooling configuration is shown in Fig. 1. The FET is mounted in an alumina heat sink.

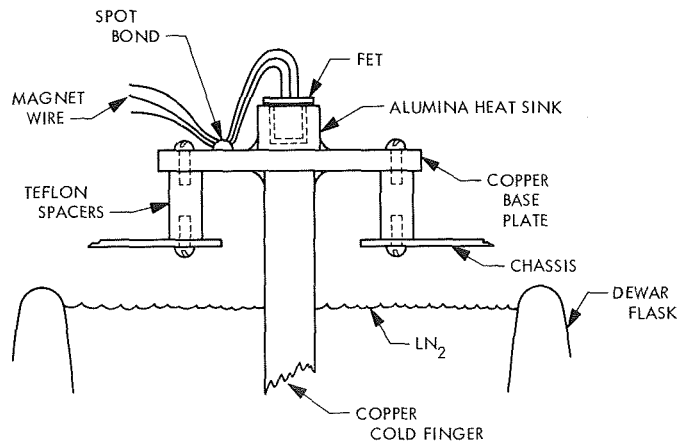


Fig. 1. Preamplifier first-stage cooling configuration

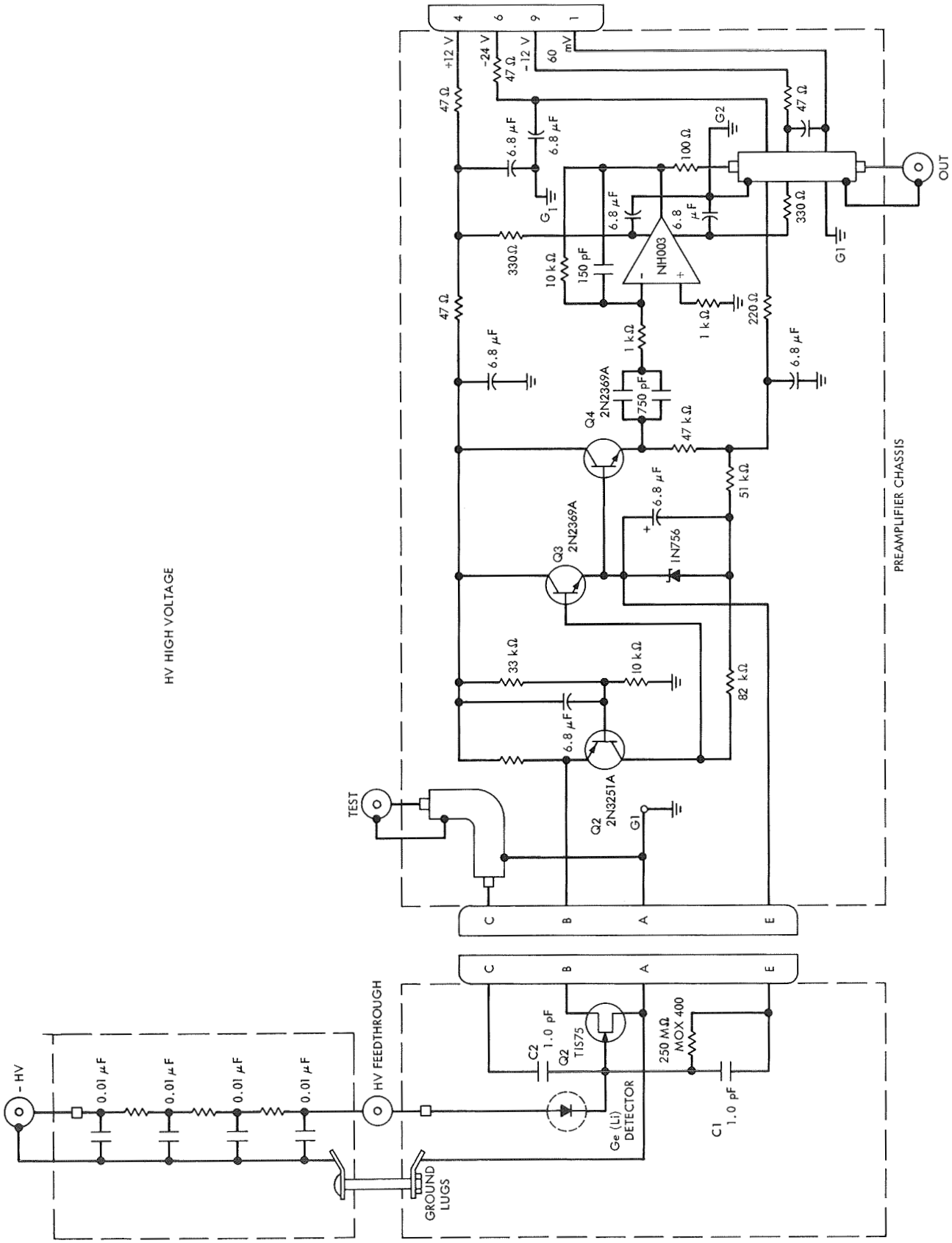


Fig. 2. Cryogenically cooled preamplifier

The heat sink provides a path of high thermal conductivity from the FET to the copper cold-finger base plate. Alumina has low electrical conductivity and, therefore, provides good insulation between the FET case and the base plate.

The FET electrical connections to the remainder of the circuit are provided using thin magnet wire to minimize heat flow through the leads. The magnet wire is spot-bonded to the base plate. The spot bond allows a large fraction of the heat in the wire to flow directly into the base plate without heating the FET. The base plate is thermally and electrically insulated from the chassis by four teflon spacers.

Base plate cooling is provided by a cold finger. The cold finger is constructed from ¼-in. copper tube, which is soldered directly to the base plate. The cold finger is inserted into the Dewar flask, which is filled with LN_2.

3. Circuitry

The circuit schematic diagram is shown in Fig. 2. The preamplifier consists of a charge sensitive amplifier followed by a shaping amplifier. The gain of the charge sensitive amplifier is determined by the integrating capacitor C1. Q1 and Q2 form a cascode voltage amplifier stage which is coupled to the output by the Darlington pair Q3–Q4. Q1 is a low-noise FET selected for optimum operation at −150°C. Q1 was selected and furnished by J. Walton, University of California Lawrence Radiation Laboratory, Berkeley. The FET resolution as a function of temperature for a 12-pF input capacity is shown in Fig. 3.

The shaping amplifier consists of an NH0003 integrated circuit with equal differentiation and integration time constants of 1.5 μs. The gain of the shaping amplifier is approximately 3 for a step function input.

4. Test Results

The cooled FET first stage consumes the major fraction of the preamplifier power. The first-stage noise at ambient temperature decreases inversely with the square root of the drain current. Test drain currents ranged from 2 to 11 mA. At ambient temperatures, the preamplifier exhibited the lowest noise at the highest drain current used in the test. During the low-temperature test, the preamplifier exhibited the lowest noise at a 5-mA drain

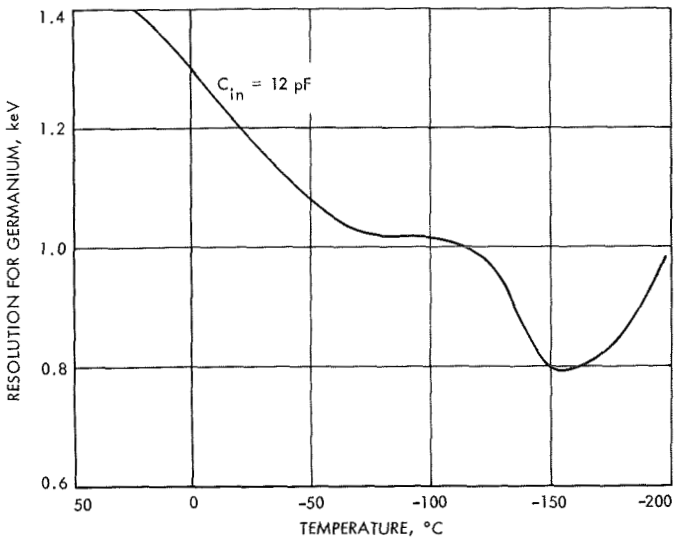


Fig. 3. Resolution as a function of temperature

current. Improved noise performance at the higher currents was offset by the heating effect on the FET resulting from the increased power dissipation.

During initial tests, some difficulty was encountered in attempting to lower the temperature of the FET to −150°C. This problem was corrected when the LN_2 level in the Dewar was raised to within a few inches of the cold-finger copper base plate. Thermocouples mounted on the cooling assembly indicated that with a cold-finger temperature of −196°C and FET dissipation of 50 mA, the baseplate temperature was −170°C and the FET case temperature was −150°C.

Bubbling of the LN_2 and acoustic noise in the test laboratory were transformed into electrical noise at the output of the preamplifier by the microphonic characteristics of the FET. The baseplate vibration was restricted by the use of sturdy teflon spacers at each corner. Vibration transmitted to the FET case by means of the connecting leads was considerably reduced by spot-bonding the leads to the base plate.

The preamplifier noise with a 30-pF input capacitance was 2.8 keV (FWHM)[1] at room temperature. The noise was measured 1.7 keV (FWHM) under cooled conditions. This noise resolution for a germainium solid-state detector represents an equivalent preamplifier input charge of 240 rms electrons.

[1]FWHM = full width at half maximum.

III. Science Data Systems

SPACE SCIENCES DIVISION

A. A Search for *r*th-Degree Tetranomials Over GF(2) With Periods of $2^{r-1} - 1$ or $2^{r-1} - 2$,

M. Perlman

1. Introduction

The (binary) linear logic feedback shift register (FSR) shown in Fig. 1 has been investigated in considerable detail (Ref. 1). The state of the ith two-state memory element at clock pulse interval (CPI) k is denoted as a_{k-i}. The behavior of the FSR can be characterized by the linear recurrence relationship

$$a_k = e + \sum_{i=1}^{r} c_i \, a_{k-i} \tag{1}$$

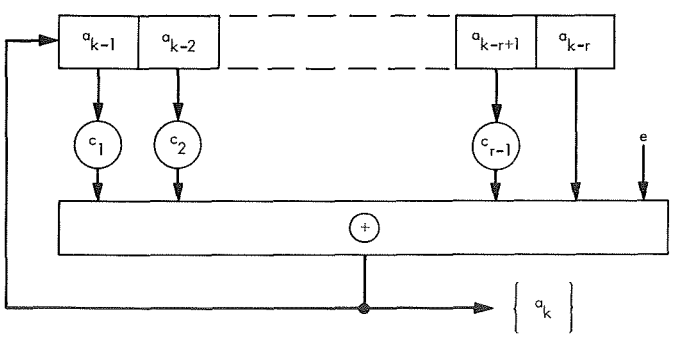


Fig. 1. *r*-Stage linear feedback shift register

The bit fed back at CPI k is denoted as a_k. The ith stage contributes to the feedback when the Boolean multiplier c_i is at state-value 1. The summations are taken modulo 2 and e, a Boolean constant, is 0 for modulo 2 summing (EXCLUSIVE-OR) or 1 for the complement of modulo 2 summing (NOT EXCLUSIVE-OR).

The cycle length or periodicity of $\{a_k\}$ can be determined from its generating function

$$G_e^r(x) = \sum_{k=0}^{r} a_k x^k$$

For $e = 0$ (Ref. 1)

$$G_0^r(x) = \frac{g_0^r(x)}{f_0^r(x)} \tag{2}$$

where

$$g_0^r(x) = \sum_{i=1}^{r} c_i x^i \left(a_{-i}x^{-i} + a_{-i+1}x^{-i+1} + \cdots + a_{-1}x^{-1}\right)$$

and

$$f_0^r(x) = 1 + \sum_{i=1}^{r} c_i x^i$$

For $e = 1$ (Ref. 2)

$$G_1^r(x) = \frac{g_1^r(x)}{f_1^r(x)} = \frac{1 + (1+x)\,g_0^r(x)}{(1+x)\,f_0^r(x)} \tag{3}$$

The characteristic polynomial $f_e^r(x)$ is a function of the feedback connections only. The least integer n for which $f_e^r(x)$ divides $x^n + 1$ is the length of the longest FSR cycle(s). This is precisely the period of $f_e^r(x)$.

The polynomial $g_e^r(x)$ is a function of the initial state of the register and the feedback connections. The degree of $g_e^r(x)$ is always less than that of $f_e^r(x)$. An initial state $a_{-1}a_{-2}\cdots a_{-r}$ which results in a $g_e^r(x)$ that is relatively prime to $f_e^r(x)$ will lie on a cycle whose length is n the period of $f_e^r(x)$. An initial state that yields a $g_e^r(x)$ that has a factor in common with $f_e^r(x)$ will lie on a cycle whose length divides n. Thus all cycle lengths of an FSR with a characteristic polynomial $f_e^r(x)$ divide n the period of $f_e^r(x)$.

The initial state $00\cdots01$ results in a $g_0^r(x) = 1$. The initial state $00\cdots0$ results in a $g_1^r(x) = 1$. Each of these initial states lies in a cycle of longest possible length where the latter is associated with $g_1^r(x)$ and the former with $g_0^r(x)$.

2. Primitive Polynomials Over GF(2)

Every irreducible polynomial of degree r over GF(2) divides $x^{2^r-1} + 1$ (Ref. 3). If $n = 2^r - 1$ is the least integer for which divisibility holds, then the irreducible polynomial is said to be primitive. A primitive characteristic polynomial $f_0^r(x)$ describes an r-stage FSR capable of generating a recurring sequence $\{a_k\}$ of period $2^r - 1$. The sequence has three properties associated with randomness (Ref. 1) and is termed a maximal-length or pseudonoise (PN) sequence.

Primitive polynomials over GF(2) exist for every degree, although complete tables have not been generated beyond degree 19, e.g., Refs. 4 and 5. A primitive polynomial over GF(2) must contain an odd number of terms, including a constant term of 1. This ensures that x or $x + 1$ are not factors. The primitive trinomial characterizes the FSR of least complexity in terms of feedback (two-tap) combinational logic. Since trinomials represent a restricted class of polynomials over a field of two elements, tables of irreducible trinomials have been generated through degree 1000 (Ref. 6). Unfortunately, there are many values of r for which an irreducible trinomial of degree r does not exist. Furthermore, there are values of r for which irreducible but not primitive trinomials of degree r exist (Refs. 1, 6, and 7).

Tetranomials of degree r over GF(2) characterize an r-stage three-tap linear FSR where

$$a_k = a_{k-i} + a_{k-j} + a_{k-r} \tag{4}$$

and

$$f_0^r(x) = 1 + x^i + x^j + x^r \tag{5}$$

Of interest are those tetranomials which contain an $(r-1)$th or $(r-2)$th degree primitive polynomial as a factor. Assume

$$f_0^r(x) = 1 + x^i + x^j + x^r = (1+x)\,\phi(x) \tag{6}$$

where $\phi(x)$ is a primitive $(r-1)$th degree polynomial. The period of the longest cycle is the least common multiple (LCM) of the periods of its irreducible factors or $2^{r-1} - 1$. Assume

$$f_0^r(x) = 1 + x^i + x^j + x^r = (1+x)^2\theta(x) \tag{7}$$

where $\theta(x)$ is a primitive $(r-2)$th degree polynomial. The period of the longest cycle is the LCM of the period of the repeated irreducible factor $(1+x)^2$ and the period of $\theta(x)$ (Ref. 2). The length of the longest cycle is

$$\mathrm{LCM}[2, 2^{r-2} - 1] = 2^{r-1} - 2$$

A search was conducted for tetranomials through degree 34 over GF(2) of the form shown in Eqs. (6) and (7).

3. Tetranomials of Degree r Over GF(2) With Periods of $2^{r-1} - 1$ or $2^{r-1} - 2$

a. Search technique. A 34-stage FSR was implemented to operate at a clock frequency of 1 MHz. Provisions were made to alter its length, insert any predetermined initial state, and feedback the modulo 2 sum (EXCLUSIVE-OR) or the complement of the modulo 2 sum (NOT EXCLUSIVE-OR) of the content of the stages i, j, and r, where r denotes the last stage and $i<j<r$. A word detector (effectively an r-input gate) was used to sense and "remember" the initial state. By means of control logic, the initial state is inserted, the word detector is primed, and clock is initiated, and the clock is terminated when the FSR is returned to its initial state.

b. Theorems. A number of theorems are useful in verifying results and reducing the number of cases to be searched.

Theorem 1 (Ref. 1)

Every state of the FSR will have a unique predecessor and a unique successor (i.e., the cycles will be branchless) if, and only if, the feedback function can be decomposed as follows:

$$a_k = F(a_{k-1}, a_{k-2}, \cdots, a_{k-r+1}) + a_{k-r}$$

where $F(a_{k-1}, a_{k-2}, \cdots, a_{k-r+1})$ is any Boolean function of the content of all stages but r. All linear feedback functions are of this form (Eq. 1).

Theorem 2 (Ref. 1)

The number of cycles for $r > 2$ into which the 2^r states of an r-stage FSR is decomposed is even or odd according to whether the number of 1's in the truth table of $F(a_{k-1}, a_{k-2}, \cdots, a_{k-r+1})$ is even or odd. When the feedback function is linear (Eq. 1), the number of cycles is even.

Theorem 3 (Ref. 1)

If $F(a_{k-1}, a_{k-2}, \cdots, a_{k-r+1}) = F(a'_{k-1}, a'_{k-2}, \cdots, a'_{k-r+1})$, states $a_{k-1}\, a_{k-2} \cdots a_{k-r}$ and $a'_{k-1}\ a'_{k-2} \cdots a'_{k-r}$ will lie on cycles of the same length which are complementary images of one another, or they will lie on the same cycle separated by 180 deg.

Table 1. Tetranomials $1 + x^i + x^j + x^r$ with periods $2^{r-1} - 1$

i	j	r
1	2	4
1	3	5
1	2	6
2	3	
1	5	7
2	4	
1	2	8
1	4	
1	6	
3	4	
2	6	9
3	5	
2	3	10
2	5	
3	6	
1	3	11
2	4	
2	7	12
1	2	14
2	5	
2	7	
2	11	
3	4	
4	9	
5	8	
3	5	15

i	j	r
5	7	15
6	8	
1	2	16
1	14	
2	11	
2	13	
3	6	
3	12	
4	5	
4	11	
5	10	
7	8	
1	11	17
4	10	
1	12	18
1	14	
2	7	
2	9	
3	8	
4	5	
4	11	
5	6	
1	7	19
6	12	
7	11	
1	14	20
2	11	

i	j	r
5	14	20
1	3	21
1	11	
3	11	
4	6	
4	12	
6	8	
8	10	
1	8	22
1	16	
2	9	
3	12	
4	7	
6	9	
1	21	23
2	8	
2	16	
1	2	24
1	12	
1	18	
2	15	
2	21	
3	8	
3	16	
3	20	
4	17	
5	12	
5	16	

i	j	r
6	7	24
6	13	
7	14	
8	9	
10	11	
2	14	25
5	19	
6	14	
8	14	
1	4	26
1	12	
1	18	
1	22	
2	11	
4	5	
4	15	
5	8	
9	10	
9	12	
9	16	
10	15	
1	5	27
1	15	
2	6	
3	7	
3	11	
5	13	
6	10	

i	j	r
7	11	27
7	17	
2	5	28
5	15	
5	22	
6	15	
6	21	
1	17	29
2	22	
3	15	
3	19	
4	24	
8	20	
11	17	
2	13	30
3	22	
4	11	
4	23	
6	7	
9	16	
11	12	
11	18	
2	12	31
9	21	
1	4	32
1	8	
1	14	
1	24	

i	j	r
1	28	32
2	19	
3	4	
3	6	
3	24	
4	15	
4	19	
4	25	
5	20	
6	7	
6	19	
7	8	
7	12	
9	16	
9	18	
10	13	
11	20	
13	14	
13	16	
2	10	33
4	14	
10	14	
11	19	
1	14	34
2	11	
2	21	
4	13	
9	14	
13	16	

Theorem 4 (Ref. 2)

A necessary but not sufficient condition for a tetranomial $1 + x^i + x^j + x^r$ to have a period $2^{r-1} - 1$ is that

$$i + j + r \equiv 1 \bmod 2 \tag{8}$$

Theorem 5 (Ref. 2)

A necessary but not sufficient condition for a tetranomial $1 + x^i + x^j + x^r$ to have a period $2^{r-1} - 2$ is that

$$i + j + r \equiv 0 \bmod 4 \tag{9}$$

The search for both types of tetranomials is further reduced by noting that a tetranomial and its reciprocal tetranomial have the same period. Also cases where i, j, or r is a multiple of 3 (say $i \equiv 0 \bmod 3$) and the remaining two integer exponents are congruent modulo 3 (i.e., $j \equiv r \bmod 3$) can be excluded since $f_0^r(x)$ would contain $1 + x + x^2$ as a factor (Ref. 1).

The results of the search appear in Tables 1 and 2 through degree 34.

Conclusions

Tetranomials of the form

$$f_0^r(x) = 1 + x^i + x^j + x^r = (1 + x)\,\phi(x)$$

with period $2^{r-1} - 1$ exist for every degree r, $4 \leq r \leq 34$, with the exception of degree 13. Each of these contains $\phi(x)$ as a factor where $\phi(x)$ is a primitive polynomial of degree $r - 1$. Thus dividing each rth degree tetranomial corresponding to an entry in Table 1 by $x + 1$ yields a *primitive polynomial* over GF(2) of degree $r - 1$. Furthermore, a maximal-length PN sequence of period $2^{r-1} - 1$ can be generated for every $r - 1$ (excluding $r - 1 = 12$) with an r-stage 3-tap FSR. An initial state of $11 \cdots 10$ yields a PN sequence associated with the characteristic polynomial $\phi(x)$. Whereas an initial state of $00 \cdots 01$ yields a complementary PN sequence associated with the characteristic polynomial $(1 + x)\,\phi(x)$ (see Theorem 3 and Example 1).

Tetranomials of the form

$$f_0^r(x) = 1 + x^i + x^j + x^r = (1 + x)^2\,\theta(x)$$

with periods $2^{r-1} - 2$ exist for every degree r, $4 \leq r \leq 34$, with the exceptions of degree 6, 8, and 14. Each of these contain $\theta(x)$ as a factor where $\theta(x)$ is a primitive polynomial of degree $r - 2$. Thus dividing each rth degree tetranomial corresponding to an entry in Table 2 by $x^2 + 1$ yields a primitive polynomial over GF(2) of degree $r - 2$. Furthermore, a sequence of length $2^{r-1} - 2$ can be generated for every $r - 1$ (excluding $r - 1 = 5, 7$, and 13). The initial state $00 \cdots 01$ lies on the cycle of length $2^{r-1} - 2$. There are two cycles which are complementary images of one another which correspond to $(1 + x)\,\theta(x)$ and $\theta(x)$, respectively. Each yields PN sequences of length $2^{r-2} - 1$. The remaining cycles yield two sequences of length 1; namely,

Table 2. Tetranomials $1 + x^i + x^j + x^r$ With Periods $2^{r-1} - 2$

i	j	r	i	j	r	i	j	r
1	3	4	5	9	18	9	15	28
1	2	5	2	3	19	1	2	29
			2	11		1	10	
1	4	7	7	10		2	17	
2	3					7	16	
			5	7	20	13	14	
1	2	9	5	11				
						5	9	30
1	6		1	2	21			
1	5	10	1	6		1	8	31
			7	12		2	27	
1	4	11				3	6	
3	6		1	13	22	3	26	
4	5		3	15		5	24	
						7	22	
1	3	12	2	11	23	9	12	
			2	19				
1	2	13	6	7		3	25	32
2	9					5	11	
4	7		1	3	24			
5	6		5	15		2	9	33
						2	25	
1	12	15	1	22	25	7	24	
2	3		3	20		10	17	
3	10		7	16				
4	5		8	15		1	13	34
						1	29	
1	7	16	1	17	26			
			3	19				
1	14	17						
3	4		1	20	27			
3	12		2	3				
6	9		9	16				

(0) and (1) corresponding to $(1 + x)$ and the sequence of length 2 corresponding to $1 + x^2$ (see Example 2).

Example 1

$$\text{Given } a_k = a_{k-1} + a_{k-2} + a_{k-4}$$

The initial state $a_{-1}\ a_{-2}\ a_{-3}\ a_{-4}$ of 0001 lies on a cycle of longest length represented by

$$G_0^4(x) = \frac{1}{1 + x + x^2 + x^4} = \frac{1}{f_0^4(x)}$$

$$= 1 + x + x^3 + x^7 + x^8 + x^{10} + \cdots.$$

The coefficients of x correspond to the bits in the recurring sequence $\{a_k\}$ whose period is 7.

k	0	1	2	3	4	5	6
a_k	1	1	0	1	0	0	0

The 16 stages of the 4-stage FSR lie on 4 disjoint cycles. They are shown in Table 3 with a_{k-i} denoted as a_i.

Table 3. 4-Stage FSR where $f_0^4(x) = 1 + x + x^2 + x^4$

k	a_1	a_2	a_3	a_4	a_k	a_1	a_2	a_3	a_4	a_k
0	0	0	0	1	1	1	1	1	0	0
1	1	0	0	0	1	0	1	1	1	0
2	1	1	0	0	0	0	0	1	1	1
3	0	1	1	0	1	1	0	0	1	0
4	1	0	1	1	0	0	1	0	0	1
5	0	1	0	1	0	1	0	1	0	1
6	0	0	1	0	0	1	1	0	1	1
0	0	0	0	0	0	1	1	1	1	1

The sequences of length 7 are PN sequences. Note that

$$f_0^4(x) = (1 + x)(1 + x^2 + x^3)$$

where $\phi(x) = 1 + x^2 + x^3$ and is primitive. The initial state 1110 yields a sequence corresponding to $1 + x^2 + x^3$. The complementary initial state 0001 yields a complementary sequence corresponding to $(1 + x)(1 + x^2 + x^3)$ or $f_0^4(x)$. The two length 1 recurring sequences (0) and (1) correspond to the $1 + x$.

Example 2

$$\text{Given } f_0^5(x) = 1 + x + x^2 + x^5$$

The six disjoint cycles are listed in Table 4.

Table 4. 5-Stage FSR where $f_0^5(x) = 1 + x + x^2 + x^5$

k	a_1	a_2	a_3	a_4	a_5	a_k	a_1	a_2	a_3	a_4	a_5	a_k
0	0	0	0	0	1	1	0	1	0	1	0	1
1	1	0	0	0	0	1	1	0	1	0	1	0
2	1	1	0	0	0	0						
3	0	1	1	0	0	1						
4	1	0	1	1	0	1						
5	1	1	0	1	1	1						
6	1	1	1	0	1	1						
7	1	1	1	1	0	0						
8	0	1	1	1	1	0						
9	0	0	1	1	1	1						
10	1	0	0	1	1	0						
11	0	1	0	0	1	0						
12	0	0	1	0	0	0						
13	0	0	0	1	0	0						
0	1	1	0	1	0	0	0	0	1	0	1	1
1	0	1	1	0	1	0	1	0	0	1	0	1
2	0	0	1	1	0	0	1	1	0	0	1	1
3	0	0	0	1	1	1	1	1	1	0	0	0
4	1	0	0	0	1	0	0	1	1	1	0	1
5	0	1	0	0	0	1	1	0	1	1	1	0
6	1	0	1	0	0	1	0	1	0	1	1	0
0	0	0	0	0	0	0	1	1	1	1	1	1

The sequences of length 7 are PN sequences. Note that

$$f_0^5(x) = 1 + x + x^2 + x^5$$

$$= (1 + x)^2 (1 + x + x^3)$$

where $\theta(x) = 1 + x + x^3$ and is primitive.

Associated with every r stage FSR corresponding to feedback configurations in Table 2 are complementary disjoint cycles of states which yield PN sequences of length $2^{r-2} - 1$ for every $r - 2$ (excluding $r - 2 = 4, 6, 12$). It may be shown that when $r > 4$ the initial state $00\cdots0101$ results in a $g_0(x) = 1 + x^2$ and lies on a cycle of length $2^{r-2} - 1$ corresponding to $\theta(x)$. The complementary state $11\cdots1010$ lies on a cycle of length $2^{r-2} - 1$ (Theorem 3) corresponding to $(1 + x)\,\theta(x)$.

A primitive trinomial does not exist of degree $r-1$ equal to 8, 12, 13, 14, 16, 19, 24, 26, 27, 30, and 32. Tetranomials of degree r do exist, however, which contain as a factor a primitive polynomial $\phi(x)$ of degree $r-1$ for each of the preceding values of $r-1$ except 12.

References

1. Golomb, S. W., *Shift Register Sequences.* Holden-Day, Inc., San Francisco, Calif., 1967.
2. Perlman M., *Pseudonoise Sequence Generation With Three-Tap Linear Feedback Shift Registers,* Technical Report No. 32-1432, Jet Propulsion Laboratory, Pasadena, Calif., November 1969.
3. Van der Warden, B. L., *Modern Algebra,* Vol. I. Frederick Ungar Publishing Co., New York, 1953.
4. Marsh, R. W., *Table of Irreducible Polynomials Over GF(2) Through Degree 19,* National Security Agency, October 24, 1957.
5. Peterson, W. W., *Error-Correcting Codes.* John Wiley & Sons, New York, 1960.
6. Zierler, N., and Brillhart, J., "On Primitive Trinomials," *Information and Control,* Vol. 13, pp. 541–554, 1968.
7. Swan, R. G., "Factorization of Polynomials Over Finite Fields," *Pac. J. Math.,* Vol. 12, pp. 1099–1106, 1962.

IV. Lunar and Planetary Sciences

SPACE SCIENCES DIVISION

A. Mars Surface Water Detection With the 400-MHz Downlink Beacon on *Viking Orbiter*, *R. J. Phillips*

1. Introduction

The possibility exists that the 400-MHz "wake-up" beacon on the *Viking Orbiter* can be used for a surface and subsurface water detection experiment.

The purpose of the experiment is to aid in the choice of a landing site (pre-separation) and for studying the spatial–temporal distribution of surface and subsurface water in the vicinity of, and at distances from, the landing site (post-separation).

The 400-MHz beacon serves as both the transmitting and receiving antenna, although an uplink receiving antenna from the lander may be used if the bandwidth, etc., is suitable. If the beacon operates in CW mode, then the Doppler-shifted surface reflection will be monitored. If the CW signal can be turned on and off to create a pulse mode, then the reflected pulse will be monitored. The pulse mode seems preferable as the returned signal can be monitored right down to zero Doppler; in the CW mode, allowance must also be made for the sidebands created by the modulating frequency. For the expected Martian slopes, if the returned signal is integrated over all Doppler frequencies, then the return should be independent of surface roughness and depend only on spacecraft position, antenna gain and power, and surface reflection coefficient. Resolution may be improved by passing the return through a series of bandpass filters, but the output of each filter will depend on rms slope as well as reflection coefficient.

The effect of dry material covering moist material has not been evaluated at this time. However, assuming a covering material with the electrical properties of the lunar regolith, cover depths up to tens of meters will not appreciably affect the ability of the experiment to detect subsurface moist rock or soil. The basic addition to the existing telemetry is a receiver–encoder.

This article discusses the results of a preliminary feasibility analysis.

2. Antenna Bandwidth and Pointing

The bandwidth on the beacon antenna is 10%, corresponding to 20 MHz on either side of the carrier or an undistorted pulsewidth of 0.025 μs. This bandwidth encompasses the anticipated Doppler shift from the surface.

It is assumed that the (cross-slotted) axis is pointing directly towards the Martian surface at periapsis. If the antenna is locked with the spacecraft in a fixed position, then, neglecting planetary curvature, the decrease in energy transmitted to the planet follows the gain function of the antenna as the spacecraft moves away from periapsis. Assuming a cosine-squared gain function, the following decibel loss versus longitude from periapsis is obtained: 0.12 dB/10 deg, 0.54 dB/20 deg, 1.25 dB/30 deg, and 2.31 dB/40 deg.

3. Dielectric Constant of Wet Versus Dry Rocks and Soils at 400 MHz

Unfortunately, there is little data available above 1×10^8 Hz on dielectric constant versus water content for rocks and soils. G. V. Keller and P. H. Licastro (Ref. 1) have measured the dielectric constant ϵ up to 1×10^8 Hz of 27 cores of various moisture content from the Morrison sandstone. It should be pointed out that rock and soil ϵ are generally much more sensitive to contained water than to composition. Extrapolating their data (as it is very smooth above 1×10^6 Hz) to 4×10^8 Hz, one obtains a straight-line fit of $\epsilon = 3.19 + 1.23 \times$ (vol% water), where all ϵ are to be taken as relative. Quadratic least-squares fits to ϵ measurements of 380 samples of rocks and soils are shown in Fig. 1 (Ref. 2). The fit to Keller's data, shown as a dashed line, appears consistent with J. H. Scott's *et al.* data. Also shown is the data of G. R. Jiracek (Ref. 3) at 1.8×10^7 Hz, which seems consistent for high-density Hawaiian basalt but inconsistent with the work done by other researchers for low-density basalt. The inconsistency, however, is within the scatter of data found by Scott *et al.*; Jiracek's measurements are single samples.

If we take the ϵ of dry surface material as 3, then for a 5 dB increase in power reflection coefficient from a homogeneous half-space, an ϵ of about 8 is required, which corresponds to 3–4% moisture (Fig. 1). An alternative approach is to consider the ϵ for dry ground uniformly distributed between 3 and 10 and the ϵ for wet ground distributed between 10 and 17. The expected difference between power reflection coefficients is then about 6.2 dB. In order for the difference between wet and dry ground to be significant, the uncertainty in other parameters must be less than, say, 2 dB. This corresponds to an uncertainty in spacecraft height of 0.26 h, where h is spacecraft altitude (e.g., 364 km at $h = 1400$ km), or a 27-deg antenna pointing error where a cosine-squared gain function is assumed.

Thus it appears that the ϵ difference between wet and dry ground at 400 MHz may be mappable.

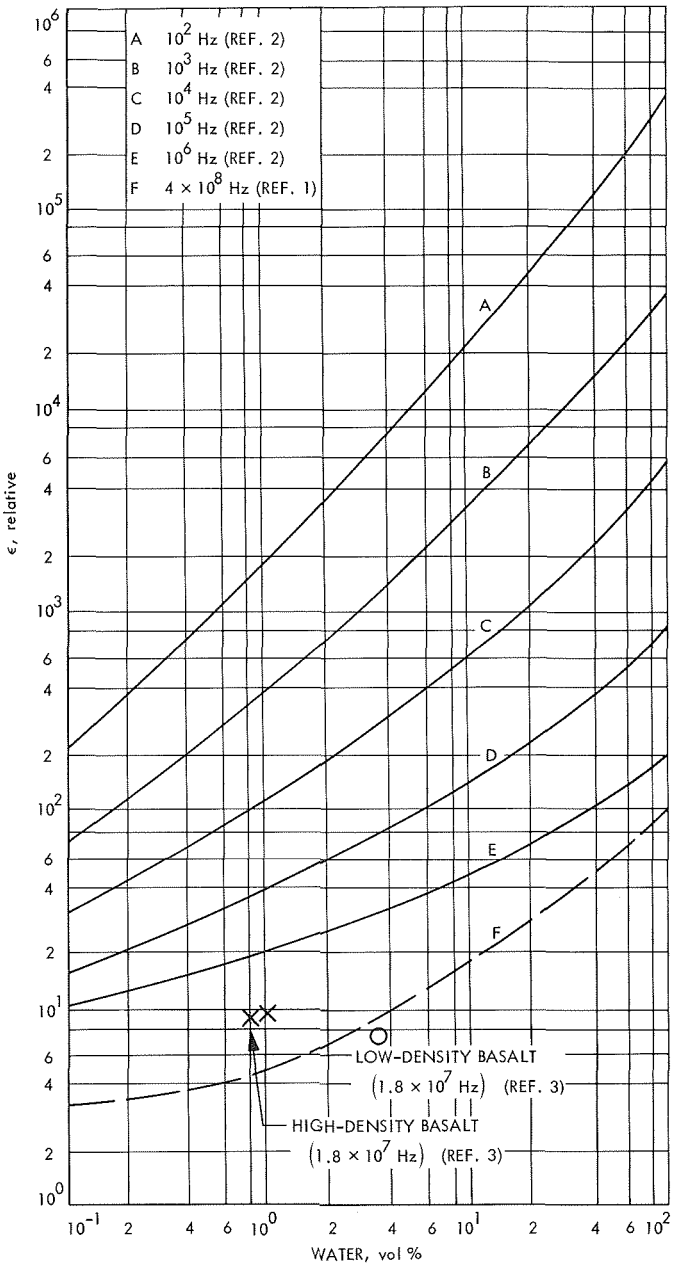


Fig. 1. Experimental values of dielectric constant versus vol% water

4. Signal-to-Noise Ratio

The power signal-to-noise ratio (SNR) has been calculated on the following assumptions: the power delivered to the beacon antenna is 2 W, the on-axis gain of the antenna is 4.5 dB and the gain function varies as cosine-squared, the rms Martian slope is 5 deg, the dominant noise source is the black-body radiation from the planet (taken at 290°K), and the bandwidth of the receiver is

2.0 kHz. For wet ground, the SNR is about 28 dB at $h =$ 1000 km and 25 dB at $h = 1400$ km. (About 5 dB can be subtracted for dry ground.) Noise in the amplifiers and filters of the receiver has not been evaluated at this time.

5. Areal Extent of Surface Return

R. B. Dyce *et al.* (Ref. 4) report that at 430 MHz, the radar return from Mars is predominantly quasi-specular with rms slopes distributed from 10 deg to less than 3 deg. Pettengill *et al.*[1] report rms slopes of about 3 deg and reflection also principally quasi-specular at 7840 MHz. Applying standard quasi-specular rough-surface scattering theory (Refs. 5 and 6), we can get some idea of the areal extent of reflection. In Fig. 2, the density and distribution functions are calculated assuming a cosine-squared gain function for the antenna, a Gaussian distribution of surface heights, a Gaussian surface correlation function, and an rms slope of 5 deg. The abscissa is horizontal distance on the ground from a point vertically beneath the antenna. The density function $f(x)$ gives the power reflected from a particular distance, while the distribution function $F(x)$ gives the integrated power return between 0 and x. Both functions have been normalized such that $F(\infty) = 1$. For rms slopes up to 15 deg, the total power returned is independent of slope. The density function peaks at a look angle normal to the one-dimensional slope (rms slope/$2^{1/2}$). It may be noted that 90% of the power returns from a circle of radius 120 km. For 3 and 10-deg rms slopes, the corresponding radius is 73 and 240 km, respectively. These numbers should be compared to the 3σ dispersion ellipse of the lander target area of 100×300 km.

[1]Pettengill, G. H., Counselman, C. C., Rainville, L. P., and Shapiro, I. I., "Radar Measurements of Martian Topography," submitted to *Astron. J.*

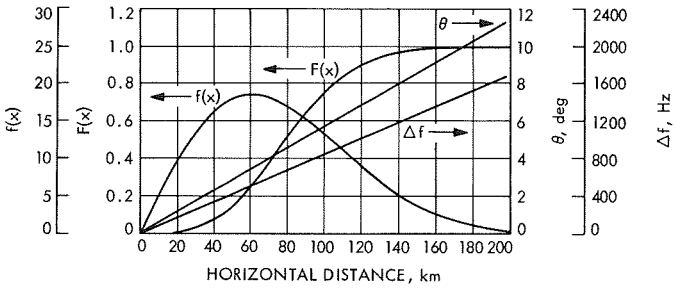


Fig. 2. Areal extent of surface return

The look angle θ shown in Fig. 2 is well within the 140-deg beamwidth of the beacon. Also shown is the Doppler shift Δf in the vicinity of periapsis for a height of 1000 km. It appears that resolution could be improved by passing the received voltage through a series of filters with bandwidths of 100 Hz.

References

1. Keller, G. V. and Licastro, P. H., *Dielectric Constant and Electrical Resistivity of Natural-State Cores,* U. S. Geological Survey Bulletin 1052-H, Washington, 1959.
2. Scott, J. H., Carroll, R. D., and Cunningham, D. R., "Dielectric Constant and Electrical Conductivity Measurements of Moist Rock: A New Laboratory Method," *J. Geophys. Res.,* Vol. 72, No. 20, pp. 5101–5115, 1967.
3. Jiracek, G. R., *Feasibility of Radio Sounding to the Groundwater Table in Hawaii,* Technical Report 10, Water Resources Research Center, University of Hawaii, 1967.
4. Dyce, R. B., Pettengill, G. H., and Sanchez, A. D., "Radar Observations of Mars and Jupiter at 70 cm," *Astron. J.,* Vol. 72, No. 6, pp. 771–777, 1967.
5. Hagfors, T., "Relationship of Geometric Optics and Autocorrelation Approaches to the Analysis of Lunar and Planetary Radar," *J. Geophys. Res.,* Vol. 71, No. 2, pp. 379–383, 1966.
6. Barrick, D. E., "Rough Surface Scattering Based on the Specular Point Theory," *IEEE Trans. Anten. Prop.,* Vol. AP-16, No. 4, pp. 449–454, 1968.

V. Physics

SPACE SCIENCES DIVISION

A. Temperature Reversal Effect in Boundary-Layer Stability Theory, *L. M. Mack*

One of the earliest predictions of the compressible boundary-layer stability theory was that cooling the wall stabilizes the boundary layer (Ref. 1). Later theoretical work (Ref. 2) considerably modified this conclusion, and experimental measurements of the effect of cooling on the transition Reynolds number have given a very mixed picture of the actual situation. In some facilities, no effect of cooling on the transition Reynolds number could be found. In others, a stabilization was found, only to be followed at lower temperatures in some instances by a destabilization. The latter effect has been termed temperature reversal. Occasionally, further cooling caused a second stabilization, an effect called re-reversal.

The theoretical picture is as follows: only the first mode is stabilized by cooling; a second-mode unstable region always exists regardless of the ratio of wall temperature to recovery temperature T_w/T_r, and in general the second-mode maximum amplification rate increases with cooling. A calculation of Reshotko (Ref. 3) based on the asymptotic theory showed the presence of a second instability zone for a certain range of temperature ratios at low Mach numbers, and so suggested that a temperature reversal and re-reversal would occur. However, a calculation of neutral-stability curves at $M_1 = 1.8$[1] directly from the parallel-flow differential equations showed the effect of cooling to be stabilizing over the complete temperature range of the calculations. Reshotko's result, therefore, appears to come from some deficiency in the asymptotic method.

The overall growth of a disturbance is more relevant to the transition problem than is the amplification rate. The growth is expressed as a ratio of the amplitude at some given Reynolds number to the initial amplitude. A calculation at $M_1 = 5.8$ (Ref. 2) showed that the effect of cooling on this quantity was susceptible to various interpretations, depending upon the definition of initial amplitude. With the initial amplitude considered to be the amplitude of a particular frequency when it first becomes unstable, cooling was found to have almost no effect on the amplitude ratio at $R = 1500$ $(R = R_x^{1/2})$. This particular amplitude ratio, A/A_1, may be thought of as the available amplification because it gives the maximum amplitude ratio that is possible for a disturbance of given

[1]Private communication with K. H. Doetsch.

frequency. On the other hand, if the initial amplitude is considered to be the amplitude at a fixed low Reynolds number, then the damping region which precedes the unstable region becomes of equal importance with the unstable region itself. The calculation showed that cooling drastically reduces this amplitude ratio, A/A_0, at $R = 1500$, and thus in this sense has a strong stabilizing influence. None of the calculations at $M_1 = 1.8$ and 5.8 from the more exact theory gave any evidence of a temperature reversal effect. On the basis of these calculations, it is only by considering the frequency response of the boundary layer to various assumed initial disturbance spectra that one can visualize how such a reversal might occur, but there is no experimental evidence that the necessary spectra exist.

The recent discovery of additional inviscid neutral solutions of the stability equations (SPS 37-56, Vol. III, pp. 71–74) now makes it possible to demonstrate a theoretical temperature reversal effect on the amplitude ratio A/A_0. It is the viscous counterparts of the additional solutions that cause the temperature reversal, and they do so by affecting the damped region which disturbances introduced at a fixed low Reynolds number must traverse before reaching the second-mode instability region. To understand this effect it is necessary to examine once more the multiple viscous solutions previously discussed in SPS 37-32, Vol. IV, pp. 222–226.

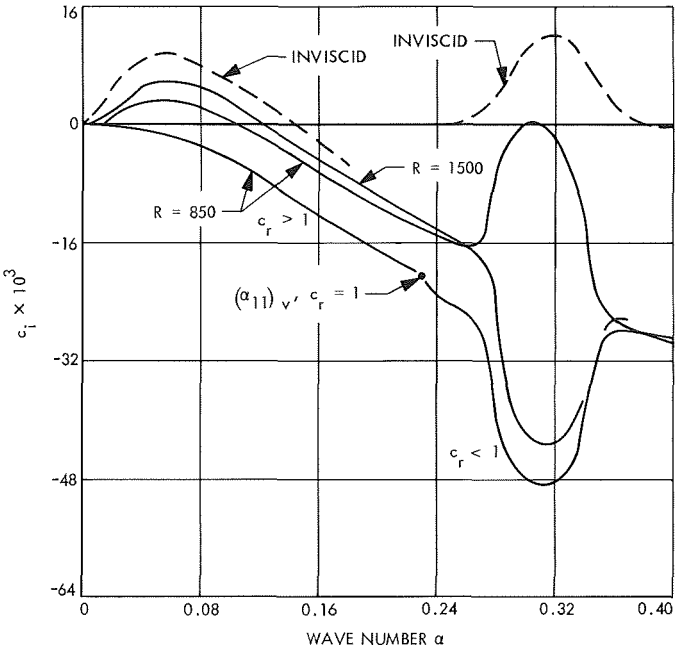


Fig. 1. Eigenvalue diagram, c_i versus α ($M_1 = 3.8$, $T_0^* = 311°K$, insulated wall)

The eigenvalues of the multiple viscous solutions at $M_1 = 3.8$ for an insulated wall are given in Figs. 1 and 2. Fig. 1 is a c_i versus α eigenvalue diagram; Fig. 2 is the dispersion relation. There are two families of solutions in the wave number range of the figures. One family starts at $\alpha = 0$, $c_0 = 1 - 1/M_1$ and is called the c_0 family. The second family is called the $(\alpha_{11})_v$ family, and was formerly thought to start at $(\alpha_{11})_v$, $c_r = 1$. The wave number $(\alpha_{11})_v$ is the viscous counterpart of α_{11}, the wave number of the inviscid regular neutral solution with $c_r = 1$. However, there are viscous damped solutions with $c_r > 1$ which connect $(\alpha_{11})_v$, $c_r = 1$ to $\alpha = 0$, $c_r = 1 + 1/M_1$, and these solutions are the viscous counterparts of the inviscid regular neutral solutions with $c_r > 1$. Thus both families of solutions start at $\alpha = 0$. At $R = 850$, the c_0 family contains both the first and second-mode amplified solutions. The $(\alpha_{11})_v$ family contains only damped solutions. At $R = 1500$, the c_0 family no longer contains the second-mode amplified solutions. It is the $(\alpha_{11})_v$ family (not shown) that contains these solutions, an arrangement that corresponds exactly to the arrangement of the multiple inviscid solutions. The Reynolds number R_i, above which the inviscid arrangement is found, is about 1300.

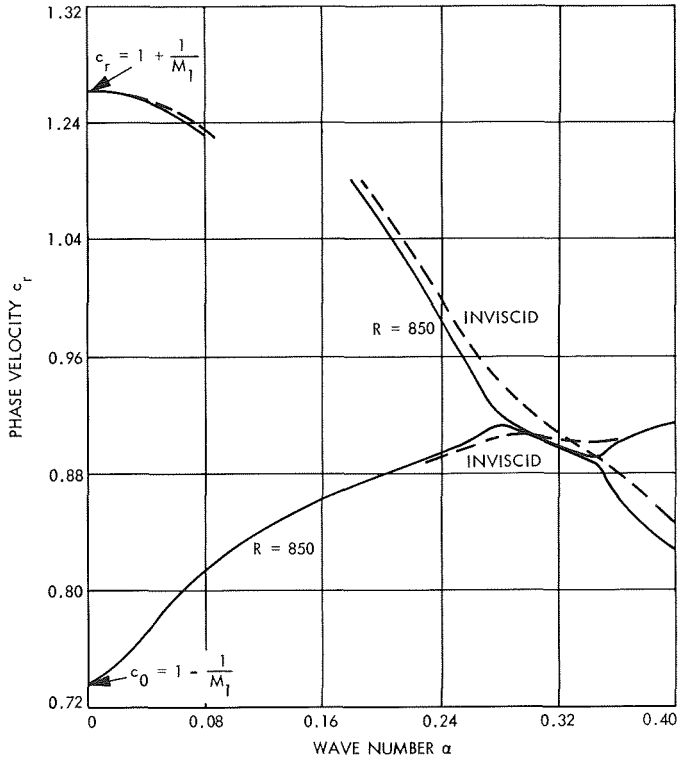


Fig. 2. Dispersion relation, c_r versus α ($M_1 = 3.8$, $T_0^* = 311°K$, insulated wall)

At $M_1 = 5.8$, R_i is large for an insulated wall. As the wall is cooled, R_i decreases. At $T_w/T_r = 0.25$, it is about 1400; at $T_w/T_r = 0.05$, it is about 400. With sufficient cooling ($T_w/T_r \leq 0.20$), there are no more amplified first-mode solutions, so the solutions of both the c_0 and $(\alpha_{11})_v$ families are damped in the first-mode region. However, the damping rate of the c_0 family, for a fixed α and R, *increases* with decreasing T_w/T_r; the damping rate of the $(\alpha_{11})_v$ family *decreases*. The question now is: To what family do the solutions belong along a constant-frequency line that leads from the origin $\alpha = 0$, $R = 0$ to the second-mode unstable region? When the second-mode amplified solutions are of the c_0 family, all solutions from $R = 0$ are also of the c_0 family; when the former solutions are of the $(\alpha_{11})_v$ family, all solutions from $R = 0$ are also of the $(\alpha_{11})_v$ family if the Reynolds number at which $\alpha = (\alpha_{11})_v$ is less than R_i. Otherwise, for $\alpha < (\alpha_{11})_v$ they are of the c_0 family. The important point is that with sufficient cooling, when an unstable second-mode solution at a given Reynolds number is traced back to the origin along a trajectory of constant frequency, all solutions will belong to the $(\alpha_{11})_v$ family rather than to the c_0 family, as at higher temperature ratios. Once this change takes place, the effect of cooling on the damping rate in the region in front of the unstable region will reverse. Instead of the damping rate increasing with decreasing T_w/T_r, as it does when the damped solutions belong to the c_0 family, it will decrease when the damped solutions belong to the $(\alpha_{11})_v$ family.

The distribution with Reynolds number of the approximate spatial amplification rate $\alpha c_i/c_r$ for two frequencies, one of which is the most unstable frequency at $R = 1500$ ($\omega\nu_1/U_1^2 = 2.0 \times 10^{-4}$), is shown in Fig. 3 at $M_1 = 5.8$, $T_1 = 125°\text{K}$, $T_w/T_r = 0.05$. Also shown for the same frequencies are the spatial amplification rates of the c_0 family of solutions. The much smaller damping of the $(\alpha_{11})_v$ family means that when the damping ceases to be governed by the c_0 family, and is instead governed by the $(\alpha_{11})_v$ family, a strong effect on A/A_0 can be anticipated. This supposition is borne out by Fig. 4, where A/A_1 and A/A_0 for the most-unstable frequency at $R = 1500$

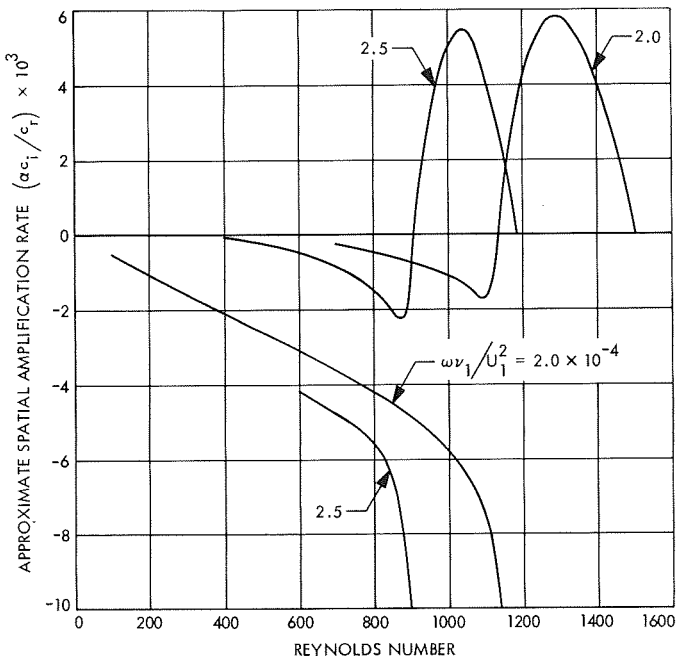


Fig. 3. Distribution of approximate spatial amplification rate with Reynolds number ($M_1 = 5.8$, $T_1^* = 125°\text{K}$, $T_w/T_r = 0.05$)

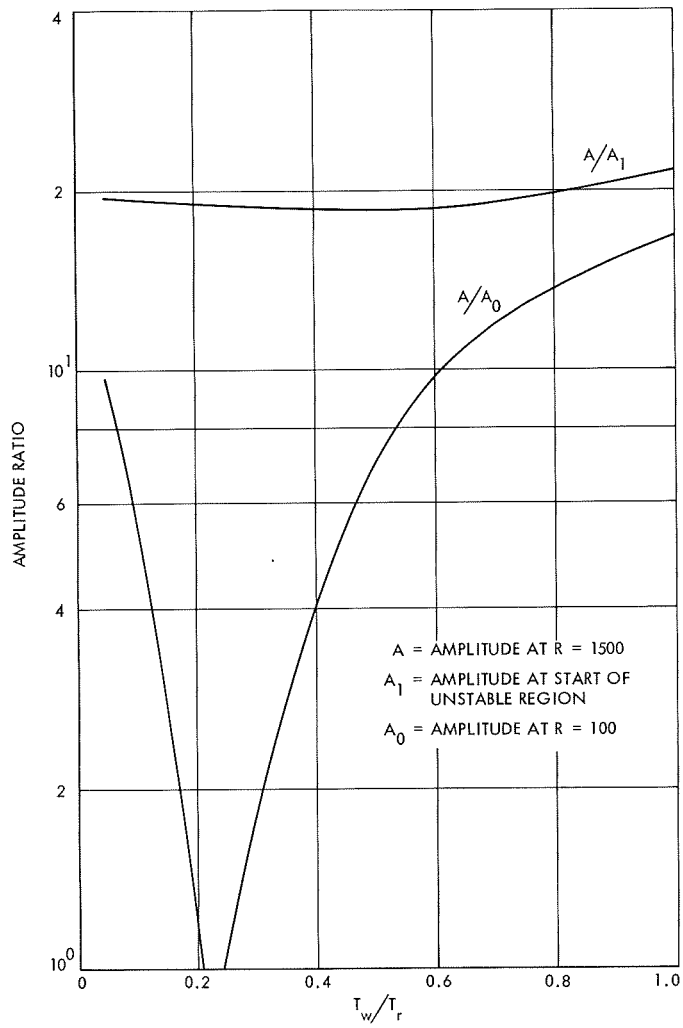


Fig. 4. Effect of cooling on maximum amplitude ratio at $R = 1500$ and $M_1 = 5.8$

are plotted against T_w/T_r. As mentioned previously, there is almost no effect of cooling on A/A_1, but there is a strong stabilizing effect on A/A_0 for $T_w/T_r > 0.24$, which comes entirely from the behavior of the damped solutions of the c_0 family. The new result is the second branch of the A/A_0 curve ($T_w/T_r < 0.21$). On this branch, which comes from the behavior of the damped solutions of the $(\alpha_{11})_v$ family, cooling has a strong destabilizing effect and A/A_0 increases with decreasing T_w/T_r. Hence, there is a temperature reversal effect on the amplitude ratio A/A_0 at low temperature ratios.

References

1. Lees, L., *The Stability of the Laminar Boundary Layer in a Compressible Fluid,* NASA Technical Report 876, National Aeronautics and Space Administration, Washington, D.C., 1947.
2. Mack, L. M., "Stability of the Compressible Laminar Boundary Layer According to a Direct Numerical Solution," *Recent Developments in Boundary Layer Research,* AGARDograph 97, Part 1, pp. 329–362, available through National Aeronautics and Space Administration, Washington, D.C.
3. Reshotko, E., "Transition Reversal and Tollmien–Schlichting Instability," *Phys. Fluids,* Vol. 6, pp. 335–342, 1963.

B. Magnetosonic Waves in the *Mariner V* Data,

T. W. J. Unti

Previous attempts, using sophisticated approaches, to subtract the effects of transverse waves in the *Mariner V* data and identify the remaining fluctuations with general magnetosonic and acoustic waves, were, in general, disappointing. A very elementary and direct method is now being pursued, which should indicate which areas of the data are most likely to yield magnetosonic waves. These areas will then be studied and analyzed with increasing degrees of sophistication.

For a magnetosonic wave, the ion density must rise with the field strength. This suggests a simple preliminary interrogation of the *Mariner V* data: The procedure is to make a statistical correlation using a computer program, calculated between density and the magnetic field component parallel to the "frozen in" field. This correlation coefficient r has been computed for 2-h intervals, for about one half of the data extending from Day 166 to Day 205 (the fast data rate portion), i.e., only those segments which appeared to have a fair chance of supporting magnetosonic waves were examined.

For a 2-h interval, r is based on 24 data points. Statistical theory can suggest which values of r may be deemed significant. Using the standard formalism as given in Ref. 1, the minimum magnitude of r which will indicate statistical correlation is given by

$$z = \frac{1}{2}\log\frac{1+r}{1-r}, \qquad \sigma_z = \frac{1}{(N-3)^{1/2}} = \frac{1}{(21)^{1/2}}$$

$$t = \frac{z - \langle z \rangle}{\sigma_z} = \frac{21^{1/2}}{2}\log\frac{1+r}{1-r}$$

where $\langle z \rangle = 0$ is equivalent to the assumption that there exists no correlation. Then

$$r = \tanh\frac{t}{21^{1/2}}$$

At the 95% significance level $t = \pm 1.96$, so that $r = \pm 0.428$. Hence $|r| > 0.43$ would imply that some correlation exists.

Of the 250 2-h intervals analyzed, only 11 had correlation coefficients exceeding $+0.5$, while 65 intervals had correlation coefficients between -0.5 and -1.0. The 11 intervals with reasonably high positive correlation are currently being further analyzed and plotted. However, the computer program also calculated parameters indicating the strength and general behavior of the field oscillations. These parameters seem to indicate that most of the 11 intervals will have the form of spikes or "boxcars" (sudden and pronounced variations in field which persist throughout most of the interval). These boxcars are thought to be a form of surface discontinuity rather than magnetosonic waves; although, if the density is positively correlated, the spike or boxcar may be propagating as a magnetosonic pulse. Two or three of the 11 intervals appear to be well behaved; detailed analysis will show whether these fluctuations satisfy the criteria of magnetosonic waves.

The fact that there are many intervals having large negative correlation coefficients seems to imply that there are many static structures in the solar wind, in which an excess of magnetic pressure on one side of the discontinuity balances an excess of plasma pressure on the other side.

For theoretical reasons, high-frequency magnetosonic waves are thought less likely to be present in the solar wind than low-frequency magnetosonic waves, and for practical reasons (the resolution for an ion measurement is 5 min) it is not expedient to search for high-frequency waves. The next direction, therefore, is to seek examples

of very-low-frequency magnetosonic waves. Running averages of the field and density measurements will be taken, and correlation coefficients will be computed.

Reference

1. Hoel, P. G., "Introduction to Mathematical Statistics," Sect. 7.1.3, John Wiley & Sons, New York, 1962.

C. Magnetic Field Line Configuration Within the Magnetosphere, *G. Atkinson and T. Unti*

In an earlier article (SPS 37-51, Vol. III, pp. 200–201) the cavity shape produced by a flow of plasma about a two-dimensional magnetic dipole field and the method of calculation was presented. This was published in more detail in Ref. 1. The results were applied to the earth's magnetosphere, and it was found that the cavity shape depended strongly on the number of field lines in the magnetic tail. As in Ref. 1, the dimensionless unit of distance in the figures will be set equal to 7.5 earth radii, and the dimensionless unit of magnetic field to 70 gammas on comparing with the magnetosphere.

The work has been extended to find the field line topology, and magnetic field strengths inside the cavity. Of particular interest is the way in which the neutral sheet distorts the field lines. Figure 5 shows field line topology with neutral sheet for $C' = 1.35$ and without neutral sheet for $C' = 0$ (Ref. 3). See SPS-51, Vol. III, for a definition of C'. X is the dimensionless distance from the magnetic dipole in the anti-solar direction. Y is the distance along the dipole axis. By appropriate matching, it is possible to obtain approximate information about the noon–midnight meridian of the three-dimensional magnetosphere. Conclusions reached from this process are:

(1) It is unlikely that the night-side auroras occur on the last closed magnetic field lines, as has been suggested. They would be expected on lower latitude field lines.

(2) Dayside auroras are associated with field lines a degree or two further south in latitude than the field line that goes to the neutral point.

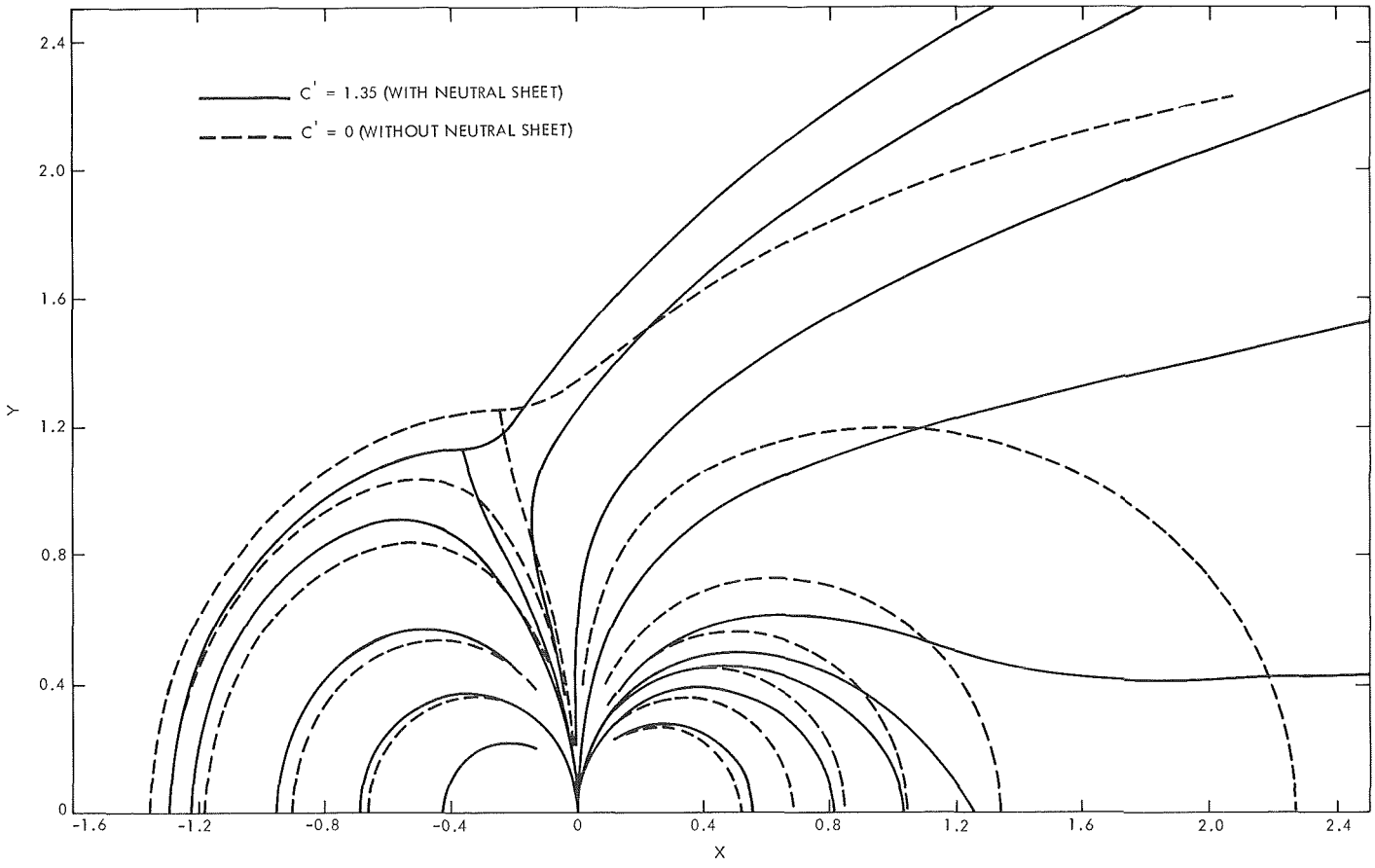


Fig. 5. Field line configurations for a typical value of tail flux and for zero tail flux

In Fig. 6, predicted field strengths are shown for the noon–midnight meridian. The predicted day–night asymmetry of 35 gammas at 6.6 earth radii is in agreement with the observations of from 30 to 60 gamma asymmetries (Ref. 3).

Figure 7 shows the predicted dropoff in magnetic field strengths down the tail, and two observational curves (Refs. 4 and 5). As expected, the three-dimensional magnetosphere tail strength drops off faster than the two-dimensional model used, and it is clear that the model is no longer representative out beyond about 40 earth radii on the night side of the magnetosphere.

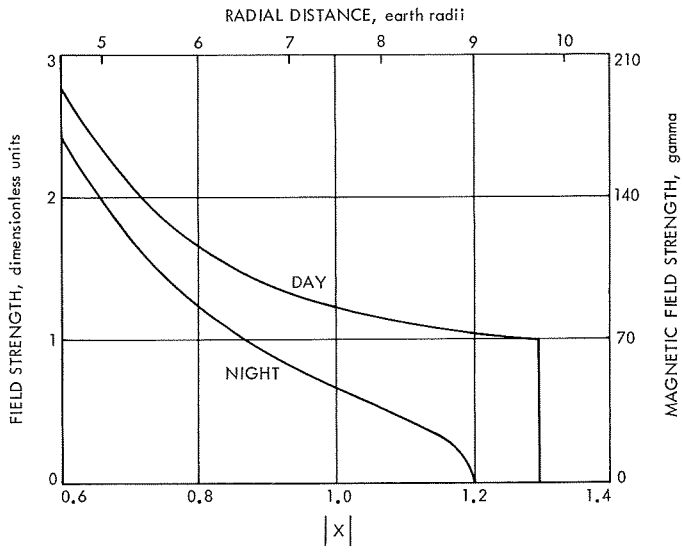


Fig. 6. Day–night asymmetry of magnetic field strength at the equatorial plane

References

1. Unti, T., and Atkinson, G., *J. Geophys. Res.*, Vol. 73, p. 7319, 1968.
2. Dungey, J. W., *J. Geophys. Res.*, Vol. 66, p. 1043, 1961.
3. Cummings, W. D., Barfield, J. N., and Coleman, P. J., Jr., *J. Geophys. Res.*, Vol. 73, p. 6687, 1968.
4. Mihalov, J. D., and Sonnett, C. P., *J. Geophys. Res.*, Vol. 73, p. 6837, 1968.
5. Behannon, K. W., *J. Geophys. Res.*, Vol. 73, p. 907, 1968.

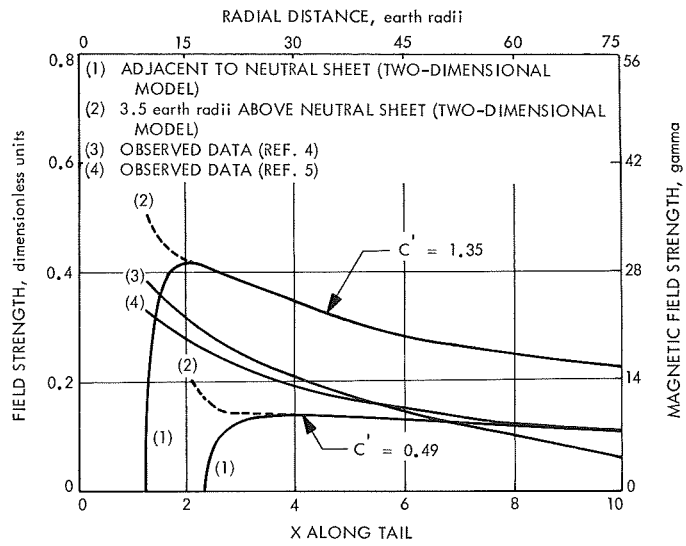


Fig. 7. Magnetic field strength in the magnetic tail

VI. Communications Systems Research

TELECOMMUNICATIONS DIVISION

A. Decoding and Synchronization Research: Binary Coding Using Single-Standard Run Lengths, *J. Molinder*

1. Introduction

A number of schemes for achieving data compression are based on redundancy reduction. That is, samples from a random source that share a common property, and thus are in some sense redundant, are in effect eliminated before transmission. Various schemes are described in Refs. 1–8. The theoretical performance limit of data compression by redundancy reduction is of course that derived by C. E. Shannon (Ref. 9). A large portion of the analysis to date has been experimental although some of the data compression schemes have recently been analyzed theoretically by assuming a specific source model (Refs. 2, 3, and 7). This is the approach followed in this article.

Under certain conditions, the amount of data required to describe a binary source may be reduced by run-length coding. Here the number of successive zeros or ones, up to some maximum run length, is transmitted rather than the individual zeros and ones. A different approach is to use $n \geq 2$ code symbols where each symbol represents a fixed run length of zeros or ones. To ensure all possible sequences can be encoded, two symbols must be used to represent a single zero and one, respectively. This leaves $n - 2$ symbols to be chosen. The technique is known as run-length coding using standard run lengths, and the problem is to choose the standard run lengths to maximize the compression ratio (CR). Experimental studies have been performed by C. Cherry, *et al.* (Ref. 1) with the "best" standard run lengths for a specific type of data being determined by numerical computations for all likely combinations of the standard run lengths.

In this article, the globally optimum single standard run lengths of zeros and ones are derived for the binary first-order Markov source. The optimum single standard run length for the binary independent source is then derived as a special case. It is assumed that the output symbols are subsequently block coded in each case. The CR maxima are found by computer search for the binary independent source when the output symbols are subsequently Huffman coded and in some cases these have been shown to be global optimums.

2. Definition of Coding Technique

A binary first-order Markov source is defined by the following transition probabilities:

$$P(0|0) = q_0 \qquad P(0|1) = p_1$$

$$P(1|0) = p_0 \qquad P(1|1) = q_1$$

where $p_0 = 1 - q_0$ and $p_1 = 1 - q_1$. This corresponds to the state diagram shown in Fig. 1.

This source is then encoded as follows:

$$\left.\begin{aligned} 0 &\rightarrow x_1 \\ 1 &\rightarrow x_2 \\ K\ 0\text{s in a row} &\rightarrow x_3 \\ N\ 1\text{s in a row} &\rightarrow x_4 \end{aligned}\right\} \qquad (1)$$

The operation of the coder may be defined by observing that no action is taken until the occurrence of one of the following events:

(1) The source changes from state 0 to state 1.

(2) The source changes from state 1 to state 0.

(3) K 0s have been accumulated.

(4) N 1s have been accumulated.

If the source changes from state 0 to state 1 (event 1), the J 0s ($1 \leq J \leq K-1$) that have been accumulated thus far are coded as J x_1s and the 1 produced by the state change is stored until it is determined whether or not $N-1$ additional 1s in a row will occur (thus allowing coding into an x_4). The source is in state 1 at the end of the coding operation. If K 0s have been accumulated (event 3), they are coded as an x_3, no input symbol is stored, and the source is in state 0 at the end of the coder operation. Similar arguments apply to events (2) and (4). Thus, the probability of a certain coder operation is dependent on whether the preceding coder operation was triggered by events (1), (2), (3), or (4). Thus, the coder operation may be defined as a mapping of certain input sequences into their corresponding output sequences as shown in Table 1. This mapping is denoted as a coder action (CA). Event (3) is equivalent to a coder output of x_3 and event (4) is equivalent to a coder output of x_4. Thus, to simplify notation, events (3) and (4) are denoted x_3 and x_4, respectively.

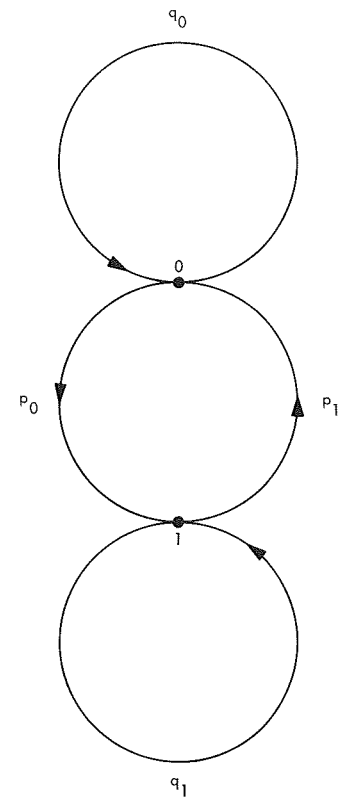


Fig. 1. State diagram

Table 1. CAs for binary first-order Markov source

Triggering event	CA	Remarks
1	$01 \rightarrow x_1$ $001 \rightarrow x_1x_1$ • • • $\underbrace{0\cdots01}_{K-1} \rightarrow \underbrace{x_1\cdots x_1}_{K-1}$	A 1 remains to be coded. The source is left in state 1.
2	$10 \rightarrow x_2$ $110 \rightarrow x_2x_2$ • • • $\underbrace{1\cdots10}_{N-1} \rightarrow \underbrace{x_2\cdots x_2}_{N-1}$	A 0 remains to be coded. The source is left in state 0.
3	$\underbrace{0\cdots00}_{K} \rightarrow x_3$	Nothing remains to be coded. The source is left in state 0.
4	$\underbrace{1\cdots11}_{N} \rightarrow x_4$	Nothing remains to be coded. The source is left in state 1.

3. Definition of CR

The CR is defined to be the expected ratio of the number of binary digits in the input sequence to the number of binary digits in the output sequence as the length of the input sequence tends to infinity. In this case, this reduces to

$$\mathrm{CR} = \lim_{n\to\infty} \frac{n}{n(x_1)l_1 + n(x_2)l_2 + n(x_3)l_3 + n(x_4)l_4} \tag{2}$$

where

n = number of input symbols

$n(x_i)$ = number of x_is ($i = 1,2,3,4$) in the output sequence

l_i = cost of the code word for x_i ($i = 1,2,3,4$) in binary digits

The optimum code is defined by the K and N that maximize the CR.

4. Derivation of CR in Terms of CAs

Referring to Fig. 1 and Table 1, and using the reasoning of *Subsection 2*, the probabilities of the possible CAs conditioned on the previous CA may be determined as follows:

$$\left.\begin{aligned}
&P_{CA}(Jx_1s \mid A) = 0 && P_{CA}(Jx_2s \mid A) = p_1 q_1^{J-1}\\
&P_{CA}(Jx_1s \mid B) = p_0 q_0^{J-1} && P_{CA}(Jx_2s \mid B) = 0\\
&P_{CA}(Jx_1s \mid x_3) = p_0 q_0^{J} && P_{CA}(Jx_2s \mid x_3) = p_0 p_1 q_1^{J-1}\\
&P_{CA}(Jx_1s \mid x_4) = p_0 p_1 q_0^{J-1} && P_{CA}(Jx_2s \mid x_4) = p_1 q_1^{J}\\
&\quad (J = 1, \cdots, K-1) && \quad (J = 1, \cdots, N-1)\\
&P_{CA}(x_3 \mid A) = 0 && P_{CA}(x_4 \mid A) = q_1^{N-1}\\
&P_{CA}(x_3 \mid B) = q_0^{K-1} && P_{CA}(x_4 \mid B) = 0\\
&P_{CA}(x_3 \mid x_3) = q_0^{K} && P_{CA}(x_4 \mid x_3) = p_0 q_1^{N-1}\\
&P_{CA}(x_3 \mid x_4) = p_1 q_0^{K-1} && P_{CA}(x_4 \mid x_4) = q_1^{N}
\end{aligned}\right\} \tag{3}$$

Thus, the conditional expectations of the number of x_1s, x_2s, x_3s, and x_4s emitted per coder action are given by

$$\left.\begin{aligned}
&E(x_1 \mid A) = 0 && E(x_2 \mid A) = \sum_{J=1}^{N-1} J p_1 q_1^{J-1}\\
&E(x_1 \mid B) = \sum_{J=1}^{K-1} J p_0 q_0^{J-1} && E(x_2 \mid B) = 0\\
&E(x_1 \mid x_3) = \sum_{J=1}^{K-1} J p_0 q_0^{J} && E(x_2 \mid x_3) = \sum_{J=1}^{N-1} J p_0 p_1 q_1^{J-1}\\
&E(x_1 \mid x_4) = \sum_{J=1}^{K-1} J p_0 p_1 q_0^{J-1} && E(x_2 \mid x_4) = \sum_{J=1}^{N-1} J p_1 q_1^{J}\\
&E(x_3 \mid A) = 0 && E(x_4 \mid A) = q_1^{N-1}\\
&E(x_3 \mid B) = q_0^{K-1} && E(x_4 \mid B) = 0\\
&E(x_3 \mid x_3) = q_0^{K} && E(x_4 \mid x_3) = p_0 q_1^{N-1}\\
&E(x_3 \mid x_4 = p_1 q_0^{K-1} && E(x_4 \mid x_4) = q_1^{N}
\end{aligned}\right\} \tag{4}$$

Since A, B, x_3, and x_4 are disjoint events whose union covers the probability space of coder actions

$$E(x_i) = E(x_i \mid A)P_{CA}(A) + E(x_i \mid B)P_{CA}(B) + E(x_i \mid x_3)P_{CA}(x_3) + E(x_i \mid x_4)P_{CA}(x_4) \quad (i = 1,2,3,4) \tag{5}$$

It can be shown (Ref. 10) that the CR converges in probability to

$$\mathrm{CR} = \frac{E(x_1) + E(x_2) + KE(x_3) + NE(x_4)}{l_1E(x_1) + l_2E(x_2) + l_3E(x_3) + l_4E(x_4)} \tag{6}$$

where $E(x_i)$ ($i = 1,2,3,4$) is as given in Eq. (5).

$P_{CA}(A)$, $P_{CA}(B)$, $P_{CA}(x_3)$, and $P_{CA}(x_4)$ must now be determined to specify $E(x_i)$ $(i = 1,2,3,4)$. This may be done by observing that the stationary probabilities of these events must satisfy the following equations: Since all probabilities refer to coder actions, the subscript CA will be dropped for notational convenience.)

$$\left.\begin{aligned}
&P(A) + P(B) + P(x_3) + P(x_4) = 1\\
&P(x_4 \mid A)\,P(A) + P(x_4 \mid B)\,P(B) + P(x_4 \mid x_3)\,P(x_3) + P(x_4 \mid x_4)\,P(x_4) = P(x_4)\\
&P(x_3 \mid A)\,P(A) + P(x_3 \mid B)\,P(B) + P(x_3 \mid x_3)\,P(x_3) + P(x_3 \mid x_4)\,P(x_4) = P(x_3)\\
&P(A \mid A)P(A) + P(A \mid B)\,P(B) + P(A \mid x_3)\,P(x_3) + P(A \mid x_4)\,P(x_4) = P(A)\\
&P(B \mid A)\,P(A) + P(B \mid B)\,P(B) + P(B \mid x_3)\,P(x_3) + P(B \mid x_4)\,P(x_4) = P(B)
\end{aligned}\right\} \tag{7}$$

Of course, these five equations are dependent since there are only four unknowns. The first four equations will be used.

$P(x_3 \mid z)$ and $P(x_4 \mid z)$ $(z = A,B,x_3,x_4)$ are given in Expression (3). Also from Expression (3)

$$P(A \mid A) = 0$$

$$P(A \mid B) = \sum_{J=1}^{K-1} p_0 q_0^{J-1} = 1 - q_0^{K-1}$$

$$P(A \mid x_3) = \sum_{J=1}^{K-1} p_0 q_0^{J} = q_0(1 - q_0^{K-1})$$

$$P(A \mid x_4) = \sum_{J=1}^{K-1} p_0 p_1 q_0^{J-1} = p_1(1 - q_0^{K-1})$$

Thus, the first four equations of Expression (7) become

$$\left.\begin{aligned}
&P(A) + P(B) + P(x_3) + P(x_4) = 1\\
&q_1^{N-1}P(A) + p_0 q_1^{N-1}P(x_3) + q_1^{N}P(x_4) = P(x_4)\\
&q_0^{K-1}P(B) + q_0^{K}P(x_3) + p_1 q_0^{K-1}P(x_4) = P(x_3)\\
&(1 - q_0^{K-1})P(B) + q_0(1 - q_0^{K-1})P(x_3)\\
&\qquad + p_1(1 - q_0^{K-1})P(x_4) = P(A)
\end{aligned}\right\} \tag{8}$$

Solving these equations and reinserting the CA notation

$$\left.\begin{aligned}
P_{CA}(A) &= \frac{(1 - q_0^{K-1})(1 - q_1^{N})}{(1 - q_0^{K}) + (1 - q_1^{N})}\\
P_{CA}(B) &= \frac{(1 - q_0^{K})(1 - q_1^{N-1})}{(1 - q_0^{K}) + (1 - q_1^{N})}\\
P_{CA}(x_3) &= \frac{q_0^{K-1}(1 - q_1^{N})}{(1 - q_0^{K}) + (1 - q_1^{N})}\\
P_{CA}(x_4) &= \frac{q_1^{N-1}(1 - q_0^{K})}{(1 - q_0^{K}) + (1 - q_1^{N})}
\end{aligned}\right\} \tag{9}$$

Note that for $K = 1$, $N = 1$, the state probabilities reduce to

$$P_{CA}(A) = 0 \qquad P_{CA}(x_3) = \frac{p_1}{p_0 + p_1}$$

$$P_{CA}(B) = 0 \qquad P_{CA}(x_4) = \frac{p_0}{p_0 + p_1}$$

where $P_{CA}(x_3)$ and $P_{CA}(x_4)$ are just the state probabilities of 0 and 1, respectively.

Using Expressions (3) and (9), and Eq. (5), the expected number of x_is emitted per coder action is given by

$$\begin{aligned}
E(x_1) &= P_{CA}(B)\sum_{J=1}^{K-1} J p_0 q_0^{J-1} + P_{CA}(x_3)\sum_{J=1}^{K-1} J p_0 q_0^{J}\\
&\quad + P_{CA}(x_4)\sum_{J=1}^{K-1} J p_0 p_1 q_0^{J-1}\\
&= p_0 \sum_{J=1}^{K-1} J q_0^{J-1}[P_{CA}(B) + q_0 P_{CA}(x_3) + p_1 P_{CA}(x_4)]\\
&= \frac{[1 - K q_0^{K-1} + (K-1)q_0^{K}](1 - q_1^{N})}{p_0[(1 - q_0^{K}) + (1 - q_1^{N})]}
\end{aligned} \tag{10}$$

Similarly,

$$E(x_2) = \frac{[1 - N q_1^{N-1} + (N-1)q_1^{N}](1 - q_0^{K})}{p_1[(1 - q_0^{K}) + (1 - q_1^{N})]} \tag{11}$$

$$E(x_3) = P_{CA}(x_3) = \frac{q_0^{K-1}(1 - q_1^{N})}{(1 - q_0^{K}) + (1 - q_1^{N})} \tag{12}$$

$$E(x_4) = P_{CA}(x_4) = \frac{q_1^{N-1}(1 - q_0^{K})}{(1 - q_0^{K}) + (1 - q_1^{N})} \tag{13}$$

In summary, the CR is given by Eq. (6) where $E(x_i)$ is given by Eqs. (10)–(13).

5. Optimal Code for Output Symbols of Equal Length

If $l_1 = l_2 = l_3 = l_4 = L$, Eq. (6) may be written

$$\mathrm{CR} = \frac{1}{L}\left[1 + \frac{(K-1)E(x_3) + (N-1)E(x_4)}{E(x_1) + E(x_2) + E(x_3) + E(x_4)}\right] \tag{14}$$

Substituting Eqs. (10)–(13) into Eq. (14) and reducing yields

$$\mathrm{CR} = \frac{1}{L}\left[1 + \frac{p_0p_1(K-1)q_0^{K-1}(1-q_1^N) + p_0p_1(N-1)q_1^{N-1}(1-q_0^K)}{p_1[1-(K-1)q_0^{K-1}+(K-2)q_0^K](1-q_1^N) + p_0[1-(N-1)q_1^{N-1}+(N-2)q_1^N](1-q_0^K)}\right] \tag{15}$$

Differentiating Eq. (15) with respect to K and setting the result equal to zero yields

$$p_0p_1\frac{q_0^{K-1}(1-q_1^N)^2(p_1+p_0)}{Y^2}[(K-1)\ln q_0 + 1 - q_0^K] = 0 \tag{16}$$

where Y is the denominator of Eq. (15).

For $K \geq 1$, $N \geq 1$,

$$p_0p_1q_0^{K-1}(1-q_1^N)^2(p_1+p_0) > 0$$

$$1-(K-1)q_0^{K-1}+(K-2)q_0^K = p_0^2\sum_{J=1}^{K-1}Jq_0^{J-1} + p_0q_0^{K-1} > 0$$

and

$$1-(N-1)q_1^{N-1}+(N-2)q_1^N = p_1^2\sum_{J=1}^{N-1}Jq_1^{J-1} + p_1q_1^{N-1} > 0$$

implying Y and $Y^2 > 0$. Thus, the left-hand side of Eq. (16) is equal to zero only if

$$(K-1)(-\ln q_0) = 1 - q_0^K \tag{17}$$

The graphical solution of this implicit equation is shown in Fig. 2.

From Fig. 3 and Eq. (16) it can be seen that if K is decreased from K_0, Eq. (16) is positive, while if K is increased from K_0, Eq. (16) is negative. This means that the slope of Eq. (16) (or, equivalently, the second derivative of Eq. 15) is negative at K_0, thus assuring that K_0 determines a maximum. It is geometrically evident from Fig. 2 that there is only one solution to Eq. (17). Thus, the integer $K = K_0$ most nearly satisfying Eq. (17) defines the globally optimum run length within ± 1.

Since Eq. (15) is symmetrical in K and N, it is clear that Eq. (15) is maximized with respect to N by choosing N to be the integer most nearly satisfying (within ± 1)

$$(N-1)(-\ln q_1) = 1 - q_1^N \tag{18}$$

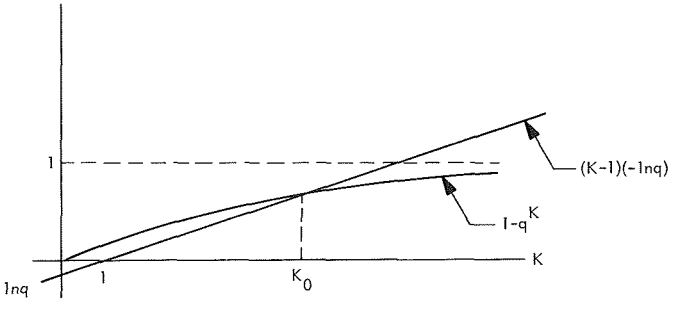


Fig. 2. Graphical solution of $(K-1)(-\ln q) = 1 - q^K$

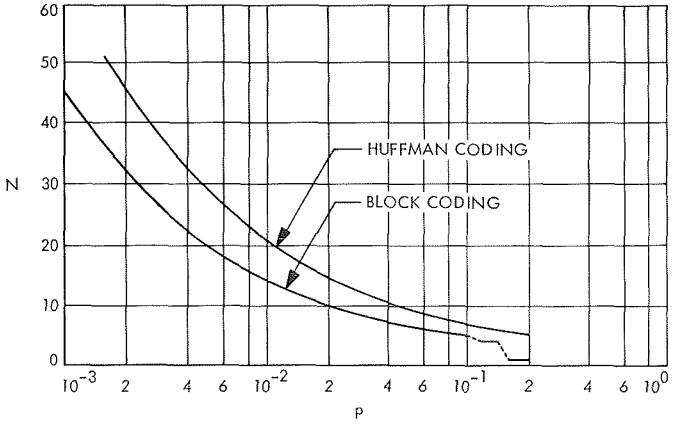


Fig. 3. Optimum N versus p for single run length

Thus, the globally optimum code is defined by the integers K and N most nearly satisfying Eqs. (17) and (18), respectively. The solutions of Eqs. (17) and (18) and the resulting CRs for various values of p_0 and p_1 are given in Table 2.

Note that in Table 2 the CR for $p_0 = 0.001$ and $p_1 = 0.500$ is greater than that for $p_0 = 0.001$ and $p_1 = 0.005$ but lower than that for $p_0 = p_1 = 0.001$. This seems strange since in the second case more strings of 1s should occur than in the first case and thus, perhaps, a greater overall CR should be expected. This behavior can be intuitively explained by the fact that for $p_0 \ll p_1$, the state probability of a zero is nearly one as shown below.

$$P(0) = \frac{p_1}{p_1 + p_0} \approx \frac{p_1}{p_1} \approx 1$$

Thus, the source is almost always in the state 0 where high CRs are obtained. As p_1 approaches p_0, the source is less likely to be in state 0 and the overall CR decreases even though the CR obtained in state 1 is increasing. Finally, as the CR in state 1 increases further, the overall CR increases again.

6. Optimal Code When Huffman Coding is Used to Code Output Symbols

The probabilities of x_i $(i = 1,2,3,4)$ may be defined as the limit of their frequency ratio as the length of the input sequence tends to infinity. Thus,

$$P(x_i) = \lim_{n\to\infty} \frac{n(x_i)}{n(x_1) + n(x_2) + n(x_3) + n(x_4)} \qquad (19)$$

where $n(x_i)$ $(i = 1,2,3,4)$ is the number of x_is in the output sequence and n is the number of binary digits in the input sequence. In terms of CAs, it can be shown (Ref. 10) that Eq. (19) converges in probability to

$$P(x_i) = \frac{E(x_i)}{E(x_1) + E(x_2) + E(x_3) + E(x_4)}$$

A finite computer search may now be performed to determine the optimum K and N for the Huffman-coded output symbols (Ref. 10).

7. Reduction to the Independent Source

If p_0 and p_1 of the binary first-order Markov source are chosen to be p and q, respectively, the Markov source is equivalent to an independent binary source with probabilities p and q for a 1 and 0, respectively.

Assuming 0 is the more likely symbol $(q \gg p)$, and that only runs of 0s are coded $(N = 1)$, Eq. (15) reduces to

$$\mathrm{CR} = \frac{1}{L}\left[1 + \frac{(K-1)pq^K}{1 + (K-1)q^{K+1} - Kq^K}\right]$$

Table 2. Optimum *K,N* and CR when output symbols are block coded (*L* = 2)

P_1 \ P_0	0.500	0.100	0.050	0.010	0.005	0.001
0.500	(1,1) 1.000	(5,2) 1.218	(7,2) 1.639	(14,2) 3.537	(20,2) 4.989	(45,2) 11.152
0.100	(2,5) 1.218	(5,5) 1.392	(7,5) 1.672	(14,5) 3.298	(20,5) 4.666	(45,5) 10.703
0.050	(2,7) 1.639	(5,7) 1.672	(7,7) 1.859	(14,7) 3.249	(20,7) 4.528	(45,7) 10.415
0.010	(2,14) 3.537	(5,14) 3.298	(7,14) 3.249	(14,14) 3.821	(20,14) 4.688	(45,14) 9.704
0.005	(2,20) 4.989	(5,20) 4.666	(7,20) 4.528	(14,20) 4.688	(20,20) 5.288	(45,20) 9.600
0.001	(2,45) 11.152	(5,45) 10.703	(7,45) 10.415	(14,45) 9.704	(20,45) 9.600	(45,45) 11.471

Key: (K,N) CR

Table 3. Optimum *N* and CR when output symbols are block coded (*L* = 2)

p	N	CR
0.5000	1	1.000
0.2000	1	1.000
0.1000	5	1.181
0.0500	7	1.636
0.0300	8	2.093
0.0200	10	2.551
0.0100	14	3.583
0.0050	20	5.046
0.0030	26	6.500
0.0020	32	7.951
0.0015	37	9.173
0.0010	45	11.224

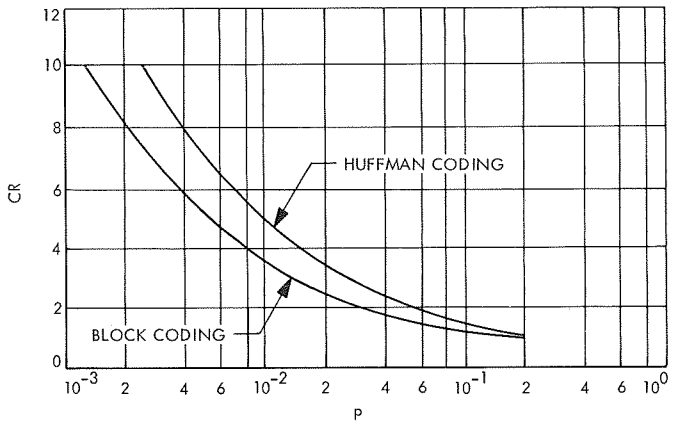


Fig. 4. CR versus p for single run length

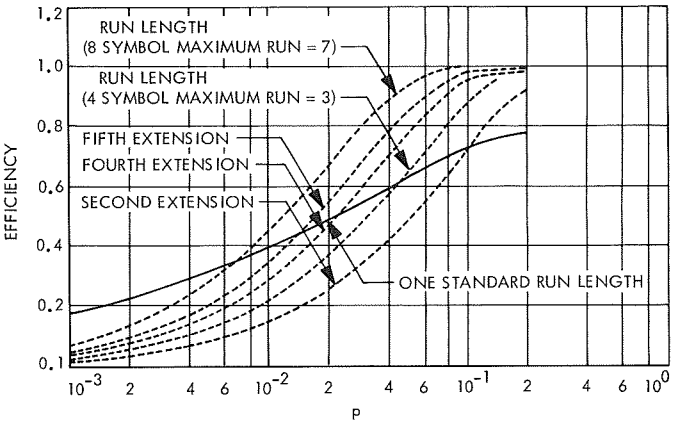


Fig. 5. Comparison of coding schemes

Table 4. Optimum N and CR when output symbols are Huffman coded

p	N	CR
0.5000	1	1.000
0.2000	5	1.102
0.1000	7	1.559[a]
0.0500	10	2.207[a]
0.0300	12	2.856[a]
0.0200	15	3.503[a]
0.0100	21	4.964[a]
0.0050	29	7.034[a]
0.0030	37	9.091[a]
0.0020	45	11.141[a]
0.0015	52	12.871[a]
0.0010	64	15.772

[a]Search carried out far enough to guarantee a global maximum.

Encoding three output symbols requires a block code length $L = 2$. The resulting CRs for various values of p are given in Table 3. Plots of the optimum N versus p and CR versus p are given in Figs. 3 and 4, respectively.

The results of the computer search when Huffman coding is used on the output symbols are given in Table 4 and plotted in Figs. 3 and 4. The points for which the search has been carried out far enough to guarantee a global maximum are identified in Table 4. A comparison of the efficiency of this scheme with various other coding schemes is given in Fig. 5 (Huffman coding is used in all cases).

References

1. Cherry, C., Kubba, M. H., Pearson, D. E., and Barton, M. P., "An Experimental Study of the Possible Bandwidth Compression of Visual Image Signals," *Proc. IEEE*, Vol. 51, pp. 1507–1517, Nov. 1963.
2. Davisson, L. O., "Data Compression Using Straight Line Interpolation," *IEEE Trans. Inform. Theory*, Vol. IT-14, pp. 390–394, May 1968.
3. Ehrman, L., "Analysis of Some Redundancy Removal Bandwidth Compression Techniques," *Proc. IEEE*, Vol. 55, pp. 278–287, Mar. 1967.
4. Gardenhire, L. W., "Redundancy Reduction—the Key to Adaptive Telemetry," proceedings of the National Telemetering Conference, pp. 1–16, 1964.
5. Kortman, C. M., "A Practical Method of Data Compression," *Proc. IEEE*, Vol. 55, pp. 253–263, Mar. 1967.
6. Robinson, A. H., and Cherry, C., "Results of a Prototype Television Bandwidth Compression Scheme," *Proc. IEEE*, Vol. 55, pp. 356–364, Mar. 1967.
7. Tunstall, B. P., "Synthesis of Noiseless Compression Codes," Research Report 67-7, School of Electrical Engineering, Georgia Institute of Technology, Oct. 1967.
8. Weber, D. R., "A Synopsis on Data Compression," proceedings of the National Telemetering Conference, pp. 9–16, 1965.
9. Shannon, C. E., and Weaver, W., *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1949.
10. Molinder, J. I., *Binary Coding Using Standard Run Lengths*, Ph.D. Thesis, California Institute of Technology, May 1969.

B. Decoding and Synchronization Research: Stochastic Theory of Generalized Tracking Systems, *D. L. Elliott*

Phase-locked loops and other tracking systems with noisy inputs have been studied for some time under the assumption that certain probability densities exist and satisfy Fokker–Planck equations. The results given in this article serve to justify this assumption mathematically.

The stochastic differential equations

$$\left.\begin{aligned} \dot{\phi} &= \Omega_0 - F_0 K[Ag(\phi) + n(t)] + \sum_{k=1}^{N} y_k, \\ \dot{y}_j &= -\frac{1}{\tau_j} y_j - \frac{1}{\tau_j}(1 - F_j)K[Ag(\phi) + n(t)], \\ j &= 1, \cdots, N \end{aligned}\right\} \quad (1)$$

are used by W. C. Lindsey (Ref. 1) in the analysis of generalized tracking systems; $n(t)$ is a gaussian white noise process with power N_0, and $g(\phi)$ is a smooth function such as $\sin\phi$.

We shall give some basic mathematical facts about such equations and the Markov processes they generate. The proofs are modifications of those in Ref. 2, which discusses a general theory of stochastic systems with smooth, globally Lipschitzian coefficients.

It has been shown by E. B. Dynkin (Ref. 3) that such a system (e.g., Eq. 1) describes a vector Markov process X with continuous sample paths. However, some additional hypothesis is needed to ensure that X has continuous transition density $p(\phi,y,t \mid \phi^0, y^0, t^0)$.

For example, suppose that $\tau_2 = \tau_1$ in Eq. (1); let

$$\xi = (1 - F_2)y_1 - (1 - F_1)y_2$$

then $\dot{\xi} = (-1/\tau)\xi$. Thus, $\xi(t)$ is deterministic, and $p(\phi,y,t \mid \phi^0, y^0, t^0)$ is singular (degenerate). A number of engineering considerations can be brought forth to show that this pathology does not occur in practical tracking loops, but it turns out that the relevant hypothesis for the general problem is *controllability* of the system when $N(t)$ is regarded as a control. By controllability of the system we mean that given any two states (ϕ^0,y^0) and (ϕ^1,y^1), for every time $T > t^0$, there exists an input $N(t)$, $t^0 \leq t \leq T$, such that $\phi(T) = \phi^1$, $y(T) = y^1$ given $\phi(0) = \phi^0$, $y(0) = y^0$.

Theorem 1

Equation (1) is controllable if all τ_i are different, $F_0 \neq 0$, and $F_j \neq 1$ for $j = 1, \cdots, N$; then the Markov process X described by Eq. (1) has a unique smooth transition density $p(\phi,y,t \mid \phi^0,y^0, t^0)$ that satisfies the Fokker–Planck partial differential equation (Eq. 12 of Ref. 1).

(See Theorem 3.5 of Ref. 2; the controllability result is in Example 12e.)

Theorem 2

For the process X, $p(\phi,y,t \mid \phi^0,y^0,t^0) > 0$ for all $t > t^0$. This is a simplified version (since the coefficient of $n(t)$ is constant) of Theorem 3.6 of Ref. 2.

If the phase angle ϕ is taken modulo 2π, then $-\pi \leq \phi \leq \pi$ in Eq. (1). In this case, we have Theorem 3.

Theorem 3

There exists an invariant probability density $p^*(\phi,y)$ for X; that is, if (ϕ^0,y^0) has density $p^*(\phi,y)$, so does $[\phi(t), y(t)]$ for $t > t^0$. As $t \to \infty$,

$$p(\phi,y,t \mid \phi^0,y^0,t^0) \to p^*(\phi,y)$$

p^* is unique, and satisfies the steady-state Fokker–Planck equation.

For proof, see Theorem 4.4 and Prop. 4.6 in Ref. 2; the proof that p^* is unique follows from the positivity of p.

References

1. Lindsey, W. C., *Nonlinear Analysis and Synthesis of Generalized Tracking Systems*, Technical Report USCEE 317, University of Southern California, Dec. 1968.
2. Elliott, D. L., *Controllable Nonlinear Systems Driven by White Noise*, Dissertation, UCLA Department of Engineering, Sept. 1969.
3. Dynkin, E. B., *Markov Processes*, Springer–Verlag, 1965.

C. Combinatorial Communication: Rearrangements of Incidence Tables, *H. Taylor*

1. Introduction

The combinatorial problem solved here is the key to the proof of the following result: For functions $f(x)$ defined on the unit interval, consider the functional

$$I(f) = \int_0^1 \int_0^1 \Psi\left[\frac{|f(x) - f(y)|}{p(|x - y|)}\right] dxdy$$

where Ψ and p are given increasing non-negative functions, with $\Psi(t)$ convex in $\log t$. Then, if f^* is the monotone

decreasing function on [0,1] equimeasureable with f (the monotonic rearrangement of f), we have

$$I(f^*) \leq I(f)$$

This inequality allows statements about the modulus of continuity of f, or the measures of sets on which f is large, to be easily shown when $I(f)$ has a known bound.[1] In particular, there are important applications to the properties of sample paths of stochastic processes.

2. Definitions

A k-table on n nodes is described as follows. Starting with an $n \times n$ table of empty squares, put a permutation of the integers from 1 to n on the diagonal from lower left to upper right. Now put dots on some of the squares above the diagonal according to the following rule: for each square say the row and column of the square intersect the diagonal at the integers b and c, and then put a dot on the square if and only if $0 < |b - c| \leq k$.

The integers on the diagonal are called nodes to carry the implication that a k-table on n nodes can be the incidence table for a certain graph on n nodes with each edge between two nodes of the graph represented by a dot in the table. It is to be understood that n and k are integers, with $0 < k \leq n$, so for fixed n,k there are exactly $n!$ k-tables on n nodes and all of them are incidence tables for the same special graph.

With fixed n,k, the particular form of a k-table on n nodes will be determined by a permutation P of the integers $1, \cdots, n$. The permutation having $1, \cdots, n$ in natural order will be called the O *permutation*. The O permutation puts dots on all the squares in the first k levels above the diagonal and leaves the remaining squares empty. Figure 6 shows the graph for $n = 11$, $k = 3$, together with two of its tables.

The following theorem, conjectured by Rodemich, will be proved here.

Theorem

For any k-table on n nodes, the squares that had dots in the O permutation can be matched in a one-to-one correspondence with dots in such a way that each dot is matched with a square it can reach by steps only to the right and/or down.

[1]Garsia, A., and Rodemich, E., *Integral Inequalities Which Control the Rate of Growth of a Function* (to be published).

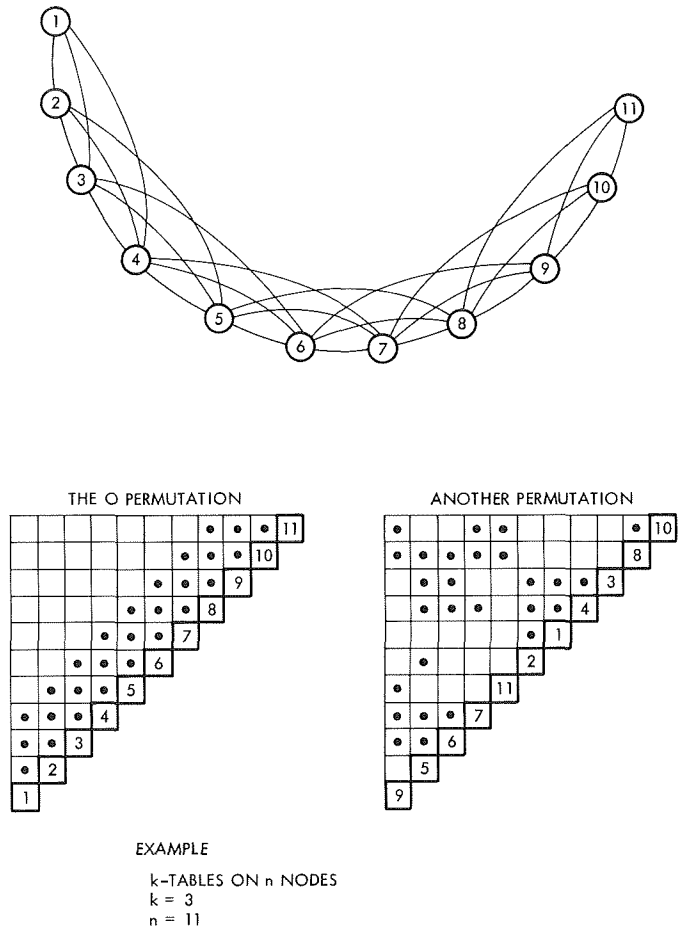


Fig. 6. k-tables on n nodes

The proof will follow from an intermediate lemma presented in *Subsection 3*.

3. An Intermediate Lemma

Let us say that an upper-left rectangle is a set of squares of a k-table determined by a square above the diagonal, together with all the squares up and/or to the left of it. Different permutations reposition nodes and dots but keep squares and rectangles fixed.

Lemma

For any k-table on n nodes, if U is any union of upper-left rectangles and P is any permutation, then

$$\#UO \leq \#UP$$

where

$\#UO$ = the number of dots covered by U in the O-permutation

$\#UP$ = the number of dots covered by U in the P-permutation

Proof

If there exist any exceptions to the lemma, there must be an exception in which the number of nodes is smallest. So suppose there is a fixed integer m for which the lemma fails, but the lemma holds for all $n < m$. This means that we have a fixed $k < m$, a union of upper-left rectangles U, and a permutation P of the k-table on m nodes such that $\#UO > \#UP$. Now we will proceed to a contradiction by showing that there must also exist a k-table on $m - 1$ nodes with a union of upper-left rectangles U' and a permutation P' such that

$$\#U'O' - \#U'P' \geq \#UO - \#UP > 0$$

as follows:

The node (m) is in line with exactly k dots in the P-permuted k-table on m nodes. Removing the node (m) from the diagonal, together with its row and column in the table, reduces the table to a k-table on $m - 1$ nodes, as can be seen simply by removing the (m) node from the corresponding graph. Now U reduces to U' and P reduces to P' in the k-table on $m - 1$ nodes. Compare the decrease in the number of dots in U with the decrease in the number of dots that ought to be in U. For this purpose, we show that

$$\#UP - \#U'P' \geq \#UO - \#U'O'$$

We now give the notation for the proof of the lemma.

O = the set of squares that had dots in the O permutation

R = the set of squares in the row to the left of the node (m)

C = the set of squares in the column above the node (m)

$$y = \#O \cap U \cap C$$
$$x = \#O \cap U \cap R$$
$$w = \#O \cap \bar{U} \cap R$$
$$v = \#O \cap \bar{U} \cap C$$
$$a = \#U' \cap O' \cap \bar{O}$$
$$d = x + y$$

Observe that

$$\#UO - d + a = \#U'O'$$

or

$$d - a = \#UO - \#U'O'$$

Figure 7 illustrates the numbers v,w,x,y,a,d for a particular case.

Since the reduction only removes squares and dots from the table without putting on any new ones, we have one general inequality,

$$\#UP - \#U'P' \geq 0$$

The node (m) has order k in the graph, and therefore $R \cup C$ has exactly k dots in the table. This gives us another general inequality,

$$\#UP - \#U'P' \geq k - (v + w).$$

The proof will be completed in two cases, which are the cases where first one and then the other of these two inequalities is helpful.

Case I $(v + w \geq k)$

In this case $U' \cap O' \cap \bar{O}$ has two disjoint subsets of size x and y respectively, hence

$$a \geq x + y = d$$

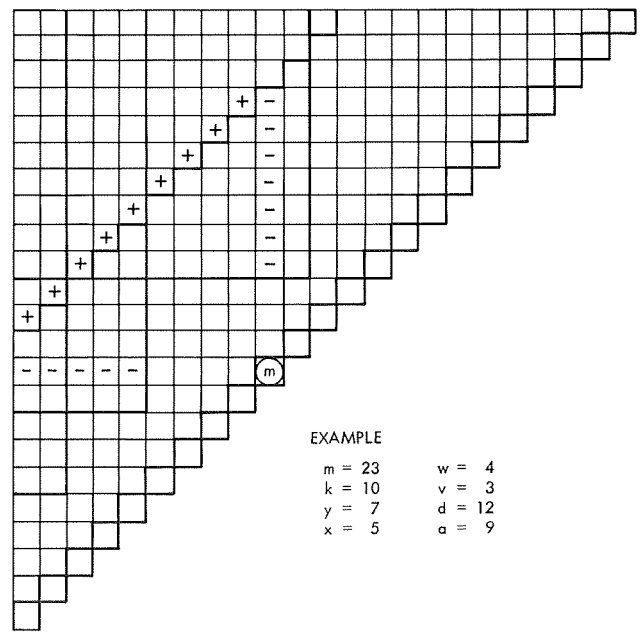


Fig. 7. Notation for the lemma

So, in this case, $d - a \leq 0$ (Fig. 8). Thus, in Case I,

$$\#UP - \#U'P' \geq 0 \geq d - a = \#UO - \#U'O'.$$

Case II $(v + w < k)$

In this case we can obtain an expression for a by means of a matching between $(R \cup C) \cap O$ and another set, as follows.

As in the example of Fig. 9, put a check ($\surd$) on all the squares of $C \cap 0$ and on all the squares of the top row which are in O and to the left of C. This makes the number of checks exactly k. Put plus signs (+) on all the squares of $O' \cap \bar{O} \cap U'$. This makes the number of plus signs exactly a.

$k > v + w$ requires that $O' \cap \bar{O} \subset U'$, and this means that every square of $R \cap O$ has either a check or a plus sign directly above it. Thus, the set of checks and plus signs is in a natural one-to-one correspondence with $(R \cup C) \cap O$. Thus

$$k + a = v + y + w + x$$

or

$$d - a = k - (v + w)$$

So, in Case II

$$\#UP - \#U'P' \geq k - (v + w) = d - a = \#UO - \#U'O'$$

Now it has been established that in either case

$$\#UP - \#U'P' \geq \#UO - \#U'O'$$

which is the same as

$$\#U'O' - \#U'P' \geq \#UO - \#UP > 0$$

An exceptional k-table on m nodes reduces to an exceptional k-table on $m - 1$ nodes in contradiction to m being the smallest number of nodes for which the lemma could fail.

That completes the proof of the lemma.

4. Proof of the Theorem

The proof of the theorem can now be done by applying the P. Hall theorem on the existence of a system of distinct representatives (Refs. 1 and 2). It is possible to

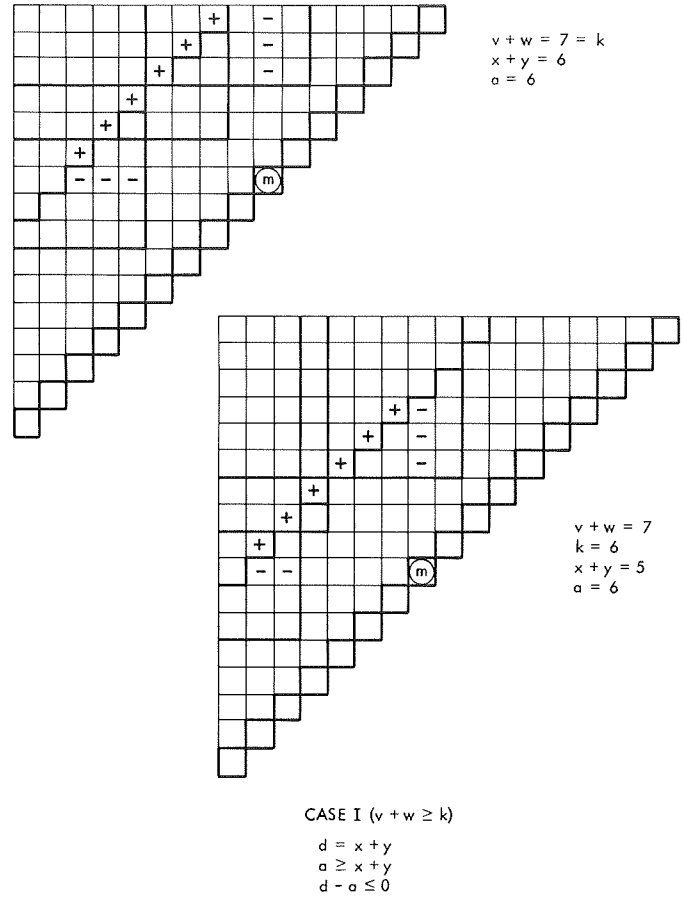


Fig. 8. Case I

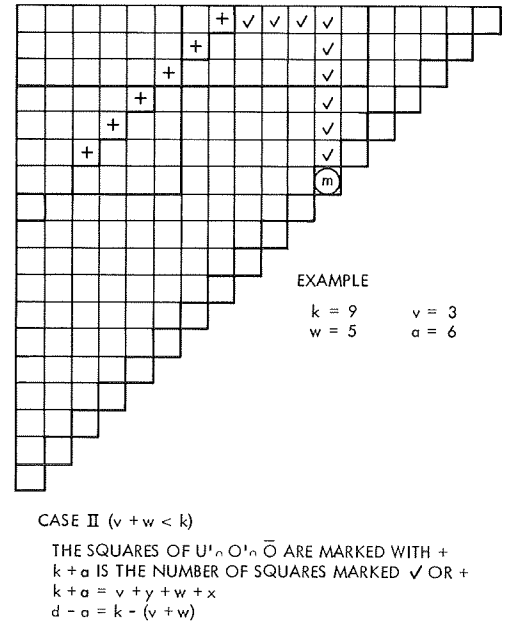


Fig. 9. Case II

choose a different element from each set in a family of sets if and only if any m fold union of sets of the family has at least m elements.

Consider the whole family of upper-left rectangles determined by squares that have dots in the O permutation. For a k-table on n nodes, this family will have

$$r = \binom{n}{2} - \binom{n-k}{2}$$

members, so they can be subscripted $S_1, \cdots, S_r$. Now in a permutation of the k-table, with the dots in new positions, we can let T_j = the set of dots on squares of S_j, for each j from 1 to r.

Stated in these terms, the theorem will be proved if we can be sure of the existence of a system of distinct representatives for the family $T_1, \cdots, T_r$.

The lemma ensures that if u is an integer with $1 \leq u \leq r$, and we have a union of u of the sets of the family $T_1, \cdots, T_r$, then there must be at least u dots belonging to the union. This is the necessary and sufficient condition for the existence of a system of distinct representatives for the family $T_1, \cdots, T_r$, according to Hall's theorem. So, for each j from 1 to r, the square that determines S_j is matched with the dot that represents T_j.

Thus, the lemma and the distinct representatives theorem give the proof of our theorem.

References

1. Hall, M., *Combinatorial Theory,* Chapter 5, Theorem 5.1.1, Blaisdell Publishing Co., 1967.
2. Ryser, H. J., "Combinatorial Mathematics," Carus Mathematical Monograph 14, Chapter 5, Theorem 1.1, *Am. Math. Assoc.*, 1963.

D. Multiple Access Communications Research: On Multiple-Demand Access Satellite Systems for Speech Communications in Remote Areas, *W. J. Hurd*

1. Introduction

The multiple-demand access communications satellite research program at JPL is concerned with investigating techniques and systems for providing reliable low-cost speech communication systems for the low-density individual user. Applications of such systems are primarily in remote areas of the world where the only existing means of communications utilize high-frequency radio, which is very unreliable. The state of Alaska, as well as the lesser-developed nations, is a prime candidate for use of such a system. For more accessible areas of the world, conventional ground-based telephone systems would probably be more economical.

It is envisioned that each system network would serve a large number of "subscribers," i.e., ground stations, but each station would be active only a small portion of the time. The research program is not concerned with service to larger communities that have local telephone networks, and which require several satellite telephone channels to other cities. Systems for such situations are discussed in Refs. 1 and 2.

This article discusses various multiple-access systems, concentrating on those system types believed most practical for the intended application. Satellite and ground-station parameters, and modulation and multiple-access methods are considered.

2. Desired System Characteristics

Some of the characteristics that a multiple-access satellite telephone communications system for individual users should have are:

(1) Low cost. The ground-station cost is especially important, since the satellite cost will be shared among many users.

(2) Reliability. The system should be reliably available for use, since servicing will be difficult in remote areas.

(3) Compatibility with different quality ground stations. The system should be usable by both economy-minded and quality-conscious users.

(4) Graceful degradation. The speech signal should be reliably intelligible with the smallest ground stations even under overload conditions.

3. Statistical Properties of Speech

The statistical properties of speech signals affect the design of a speech communication system in two manners. First, the statistical properties of the speech waveform, such as dynamic range and spectrum, determine some of the characteristics required of the channel. Second, the required perceived quality and intelligibility of the received speech signals places restrictions on the allowable distortions of the speech signal, primarily on

the signal-to-noise ratio (SNR). The various assumptions about speech signals that will be used throughout this article are given in the remainder of this subsection.

a. Waveform properties

Bandwidth. The spectrum of speech signals extends from near zero frequency to about 8 kHz. However, the spectrum falls off quite rapidly above about 1 kHz. For high-quality telephone, the transmission bandwidth is typically from 300 Hz to 3.4 kHz. This entire bandwidth is not required, however, and we will assume here a bandwidth of from 300 Hz to 3 kHz.

On–off properties of telephone signals. Since a telephone channel is considered to be a one-way channel, two channels are required for a normal conversation. A channel is considered to be in use whenever its user is either talking or listening, and the user is said to be an *active user*. There are, therefore, normally twice as many active users as conversations. On the other hand, to conform to standard telephone nomenclature, a *channel* is considered to be active only when a speech signal is present on that channel, not when that user is listening or during conversational pauses. While in use, a channel is active only 25–40% of the time; we conservatively use the 40% figure.

Amplitude characteristics. Speech signals have a very wide dynamic range of perhaps 30 dB between loud vowels and soft consonants. In a peak power limited system, it is important to keep the average power as high as possible without clipping the speech too much. When the speaker makes some attempt to control his speech volume, the average power of an active channel can be maintained at about 12–13 dB below the clipping level without causing the peaks to be clipped more than 1% of the time. We will assume the average power of an active channel to be 10 dB below the average power of a maximum-amplitude sine wave.

b. Speech quality and intelligibility.

Speech signals are almost completely intelligible under very noisy and distorted conditions because of the redundancies in speech. Significantly higher quality is required, however, before typical listeners will judge the quality as acceptable or good for most applications. Unfortunately, because of subjective differences and differences in measurement techniques and types of speech material, it is difficult or impossible to accurately determine the SNRs required for various intelligibilities or qualities.

Measurement techniques. The most common method of measuring the level of a speech signal is with a volume-unit (vu)-meter, which is a specially designed instrument that measures speech "volume" in vu. Measurements on vu-meters vary a decibel or so for the same signal, depending on the observer, and can only approximately be converted to average power, which is a more useful statistic for engineering analysis. However, studies of speech intelligibility and quality generally use vu readings, so in relating these factors to average power, we will tacitly use the conversation formula (Ref. 3)

$$P = V + 0.115\,\sigma^2 - 1.4\ \text{dBmW}$$

(Definitions of symbols used in this article are given in Table 5.)

Quality and intelligibility. We will assume the qualities and intelligibilities versus speech channel SNR given in Table 6. The SNR is the ratio of average speech power to average noise power in the frequency band from 300 Hz to 3 kHz. It should be realized that the accuracy of the table is subject to the judgment of each individual listener and speaker and can vary by several decibels. One can conservatively state, however, that the monosyllabic word intelligibility is at least 90% when the SNR is 10 dB (Refs. 4–7). This corresponds to almost perfect sentence intelligibility.

Table 5. Nomenclature

$A_{eff} = 0.55\,\pi D^2/4$	= effective area of ground station antenna
B	= speech bandwidth, Hz
D	= diameter of ground station antenna, ft
G_S	= satellite antenna gain, dB
N	= number of channels in use
$N_0 = kT$	= receiver noise spectral density
P	= average speech power for an average-talker active channel
P_G	= maximum ground station sine-wave power, dBW
P_r	= average received power per active channel
P_{sat}	= maximum satellite sine-wave power
$P'_{sat} = P_{sat}/N$	= satellite power per channel in use
P_T	= average transmitter power, dBW
SNR_{IF}	= SNR in Carson's rule pre-detection filter bandwidth
T	= receiving system noise temperature, °K
V	= average vu-meter reading
d	= slant range
k	= Boltzmann's constant
β	= modulation index
σ^2	= variance of vu-meter readings

Table 6. Ratings of speech quality versus SNR

SNR, dB	Rating
< 5	Not intelligible
10	Intelligibility good, quality poor, not acceptable (greater than 90% monosyllabic word intelligibility)
13–16	Intelligibility excellent, fair quality, acceptable for low-cost station
20	Quality good
25	Quality excellent
> 35	Extremely high quality, satisfactory for satellite trunk lines according to rigid CCIR specifications

4. Fundamental Tradeoffs Between System Types

The major current application of multiple-access satellites to telephony is in providing high-quality trunk line service to a few large ground terminals with several telephone channels per ground station. The most commonly proposed system types for this application are wideband frequency modulation–frequency-division multiple access (FM–FDMA) and pulse-code modulation–time-division multiple access (PCM–TDMA). For lower quality systems, amplitude modulation (AM) or narrowband FM can be used with FDMA. Code-division multiple access (CDMA) can also be used with either AM or FM.

a. Wideband FM and PCM. The major advantage of wideband FM–FDMA and PCM–TDMA is that they are efficient in their use of satellite power when high quality is required. SNR enhancement is attained through bandwidth expansion. Both, however, suffer three major disadvantages: (1) wide bandwidths are required per speech channel—about 80–100 kHz per channel for PCM–TDMA and 160–265 kHz per channel for FM–FDMA in order to meet CCIR (International Radio Consultative Committee) quality specifications in 60- to 240-channel systems (Ref. 8); (2) access is not random so that operator control is required to assign channels on demand; and (3) degradation is not graceful because of both the fixed number of channels limited by bandwidth and thresholding effects that cause the speech signal SNR to degrade more rapidly than the RF SNR at low SNRs. PCM–TDMA has the additional disadvantage of requiring system timing. This is not a major disadvantage when large, complex ground stations are acceptable, but is a severe limitation for small, inexpensive ground terminals.

b. AM and low modulation index FM. In a satellite system for the individual user, significantly lower quality is required than in a trunk line system. Indeed, the main reason that the CCIR specifies such high quality for satellite circuits is that the satellite link is only one portion of an overall circuit, and it is desirable that the satellite circuit noise not contribute significantly to the overall circuit noise. AM and wideband FM with fairly low modulation indexes, say $\beta < 5$, have significant advantages over high β wideband FM and PCM in the small ground station, individual user environment. Some of these advantages are as follows:

(1) Good quality is achievable without complex ground stations and with reasonable satellite power.

(2) Since less bandwidth is required per channel address, random-access addressing can be achieved within a reasonable overall system bandwidth. This eliminates the need for system bookkeeping and channel switching, which simplifies the overall system and reduces the cost.

(3) Low β FM and AM systems have graceful degradation under overload and low RF SNR conditions, whereas PCM and high β wideband FM are subject to severe thresholding. This is especially important in a system having several different quality ground stations. It is important that intelligible communication be possible with even the smallest ground stations under severe overload conditions.

The statistical properties of speech tend to favor suppressed-carrier AM in the downlink, which is average-power limited. This is because of the high peak-to-average power ratio in speech signals. There is no similar advantage in the uplink, which is peak-power limited.

c. CDMA. Spread-spectrum CDMA has also been considered. CDMA can be either orthogonal or non-orthogonal and either AM or FM can be used. The carriers are a set of codewords, often pseudonoise (PN) sequences. For AM, the speech signal is multiplied by the carrier codeword, and then modulated onto the RF carrier. For FM, the speech signal frequency modulates an FM carrier and then this signal is multiplied by the codeword.

Orthogonal CDMA. A CDMA system is orthogonal and, hence, free of interchannel interference when the codeword subcarriers are orthogonal, i.e., when the pairwise correlations of the codewords are all zero. The performance of such a system is the same as that of the corresponding AM or FM–FDMA system. One way to

closely approximate an orthogonal system is to use the distinct phases of one maximal length linear shift register pseudonoise (PN) sequence as codewords. The cross correlations between codewords will be essentially zero. Addressing would be simpler than for FDMA, because there would be no need to generate a large number of carrier frequencies. The minimum bandwidth occupancy for an orthogonal CDMA system is the same as that for double-sideband AM–FDMA if AM is used, or twice as great as FM–FDMA if FM is used. The major disadvantage to orthogonal CDMA is the need to provide accurate system timing.

Non-orthogonal CDMA. A CDMA system is non-orthogonal when the cross correlations of the codewords are non-zero. Overall system timing is not required, but the receiving station must be able to phase lock on the appropriate codeword, i.e., achieve frame synchronization and subcarrier identification. Interchannel interference cannot be avoided, and this interference is small only when large system bandwidths are used. Non-orthogonal CDMA is advantageous mainly in secure systems, where the bandwidth expansion is used for security and low quality due to interference is acceptable. Its use will not be considered further in this article. The remainder of this article will consider only AM and low β FM, with either FDMA or orthogonal CDMA.

d. Flux density limitation. By international agreement, there has been established a maximum flux density that may be caused at the earth's surface by a satellite system. This limitation is (Ref. 9): "That, in frequency bands in the range 1 to 10 GHz shared between communication-satellite systems and line-of-sight radio-relay systems, the maximum power flux-density produced at the surface of the earth by emissions from a space station, for all conditions and methods of modulation, should not exceed $[-152 + (\theta/15)]$ dB rel. 1 W/m^2 in any 4-kHz band where θ is the angle of arrival of the wave (degrees above the horizontal)" under free-space propagation conditions.

AM and narrowband FM–FDMA systems could well exceed this limitation, especially when low-quality ground stations are served. We will assume here that this problem, if it exists, will be overcome either by using a non-shared spectral band or by some simple technique of spectral smearing by modulation introduced at the satellite and removed at the ground stations.

5. SNRs

The speech channel SNRs for AM and FM systems can be expressed in terms of the ratio P_r/N_0. We first determine this ratio as a function of the system parameters, using a perhaps typical ground station as an example. Then, this ratio is related to the speech channel SNR for AM and FM systems.

The ratio P_r/N_0 can be expressed in terms of G_S, D, and P_T. In these terms, the ratio is independent of frequency and is given by

$$\frac{P_r}{N_0} = \frac{P_T G_S A_{\text{eff}}}{kT4\pi d^2} = \frac{P_T G_S (0.55)\pi D^2/4}{kT4\pi d^2}$$

a. Example system. Let us now make some reasonable assumptions as to a possibly typical system application and evaluate P_r/N_0. We assume a synchronous satellite with G_S fixed at 30 dB; the antenna then illuminates an area of the earth about 2000 mi in diameter. We also assume a $D = 10$ ft, and $T = 600$ and 1200°K for the downlinks and uplinks, respectively. Then, at the maximum d of 41,200 km (5 deg elevation) and allowing 5 dB degradation for G_S variation and rain margin, the resultant P_r/N_0 is at least $P_T + 65.5$ dB for the uplink and at least $P_T + 68.5$ dB for the downlink.

To determine P_r/N_0 for other system parameters, we note that the ratio is proportional to G_S, D^2 and $1/T$, so

$$P_r/N_0 = P_T + (G_S - 30) + 10\log_{10}\left[\left(\frac{D^2}{10^2}\right)\left(\frac{600}{T}\right)\right] + 68.5 \text{ dB}$$

Downlink. We have seen in *Subsection 3-a* that speech is present on a channel at most 40% of the time when the channel is in use. This results in a 4 dB gain in the downlink P_r/N_0. On the other hand, the signal at the satellite is essentially gaussian rather than sinusoidal, so, to avoid excessive clipping, the average satellite power must be restricted to about 7 dB less than the maximum linear sine-wave power.

One other source of degradation in downlink power is the imbalance between uplink received powers. We will assume that the uplink received power for each channel can be maintained within 2 dB of the average channel uplink received power. The net effect of the channel activity factor, the reduced transmitter power for gaussian signal, and the channel imbalance is a 5 dB reduction in effective transmitter power. Therefore, the minimum downlink P_r/N_0 for the example system is

$$\left(\frac{P_r}{N_0}\right)_{\text{down}} = P_{\text{sat}} - 10\log_{10}N + 63.5\ \text{dB}$$
$$= P'_{\text{sat}} + 63.5\ \text{dB}$$

Uplink. For an FM system, the uplink P_T while active is equal to P_G. The average uplink SNR for the example system is, then,

$$\left(\frac{P_r}{N_0}\right)_{\text{up, FM}} = P_G + 65.5\ \text{dB}$$

For suppressed-carrier AM, the average speech power is about 10 dB below P_G, so

$$\left(\frac{P_r}{N_0}\right)_{\text{up, AM}} = P_G + 55.5\ \text{dB}$$

b. Speech channel SNR for AM. AM receivers can be either coherent or non-coherent, with coherent receivers being more complex but performing 3 dB better. We will assume here that the additional complexity is worthwhile in this power-critical environment, so that coherent suppressed-carrier AM is assumed. The speech channel SNR is then $(P_r/N_0)/B$. We will assume throughout that the speech channel is from 300 Hz to 3 kHz, so $B = 2700$ Hz. The minimum AM speech channel SNR for the downlink of the example system is then

$$\text{SNR}_{\text{down, AM}} = \left(\frac{P_r}{N_0}\right)_{\text{down}} \frac{1}{B} = P'_{\text{sat}} + 29.2\ \text{dB}$$

For the uplink, the minimum SNR is

$$\text{SNR}_{\text{up, AM}} = \left(\frac{P_r}{N_0}\right)_{\text{up, AM}} \frac{1}{B} = P_G + 21.2\ \text{dB}$$

Speech channel SNR for FM. The SNR for FM depends on β, on preemphasis gains, and on P_r/N_0. Performance, however, is also subject to thresholding effects. Above threshold, the SNR for a maximum-amplitude linear sine wave with no preemphasis is

$$\text{SNR}_{\text{FM, sine}} = \frac{3P_r}{2N_0}\,\frac{\beta^2 B_2^2}{(B_2^3 - B_1^3)}$$

where $B_2 = 3$ kHz is the maximum speech frequency and $B_1 = 300$ Hz is the minimum speech frequency (Ref. 10). Since $B_1^3 \ll B_2^3$, this expression reduces to

$$\text{SNR}_{\text{FM, sine}} = \frac{3}{2}\beta^2\frac{P_r}{N_0 B_2}$$

The above expression is valid only when the SNR in the FM signal bandwidth exceeds a threshold value. This value ranges from about 4 dB for small β to about 10 dB for large β and can be reduced by techniques called *threshold extension* (Refs. 2 and 9). By Carson's rule (Refs. 9 and 10), the required FM bandwidth is $2(1+\beta)B_2$, so the SNR in the FM bandwidth is

$$\text{SNR}_{\text{IF}} = \frac{P_r}{2(1+\beta)B_2 N_0}$$

For speech signals, the average speech power level is about 10 dB below the level of a maximum amplitude sine wave. Some of this loss can be regained, however, by use of preemphasis. Assuming a 4-dB preemphasis gain, there is a net 6-dB loss relative to sine-wave signals with no preemphasis, so

$$\text{SNR}_{\text{FM}} = \frac{1}{4}\,\frac{3}{2}\,\beta^2\,\frac{P_r}{N_0 B_2}$$

For the example system, the minimum downlink SNR is

$$\text{SNR}_{\text{down, FM}} = P'_{\text{sat}} + 10\log_{10}\beta^2 + 24.5\ \text{dB}$$

We see that FM is better than AM for the downlink whenever

$$10\log_{10}\beta^2 > 4.7\ \text{dB}$$

or when β exceeds 1.72. This is, of course, subject to the received power exceeding the FM threshold.

For the uplink, the minimum FM SNR for the example system is

$$\text{SNR}_{\text{up, FM}} = P_G + 10\log_{10}\beta^2 + 26.5\ \text{dB}$$

It is important to note that, for the same β, the advantage of FM over AM is 10 dB greater in the uplink than in the downlink. This is due to the peak-to-average power ratio of speech, and because the uplink is peak-power limited whereas the downlink is average-power limited. For fixed parameters,

$$\frac{\text{SNR}_{\text{up, FM}}}{\text{SNR}_{\text{up, AM}}} = 10\,\frac{\text{SNR}_{\text{down, FM}}}{\text{SNR}_{\text{down, AM}}}$$

6. Comparison of AM and FM

In this subsection, we compare the relative merits of AM and FM by considering SNRs, bandwidth, and other factors.

a. SNR comparisons. The uplink and downlink speech channel SNRs are compared for AM and FM systems in Figs. 10 and 11. The abscissas are in units of transmitter powers for the example system; however, other system parameters would cause identical translations in the axes for all curves so that the comparison would remain valid.

For the downlink (Fig. 10), there is no clear-cut choice of either AM or FM on the basis of the SNR comparisons versus satellite power when the minimum acceptable speech channel SNR is low. FM is clearly superior, however, if good quality speech is required.

Let us assume that SNR_{IF} must exceed 4 dB in order to exceed the FM threshold. This threshold level can be obtained by use of a threshold extension demodulator [Refs. 2 (Annex 3A) and 9]. The maximum FM modulation index that can be used now depends on the system design point, i.e., on the speech channel SNR at threshold. If the threshold SNR is chosen to be 14 dB, then the maximum β is about 2.1, and the required satellite power per active user is -16.8 dBW. With AM, the 14-dB SNR is achieved with $P'_{sat} = -15.2$ dBW, so the maximum

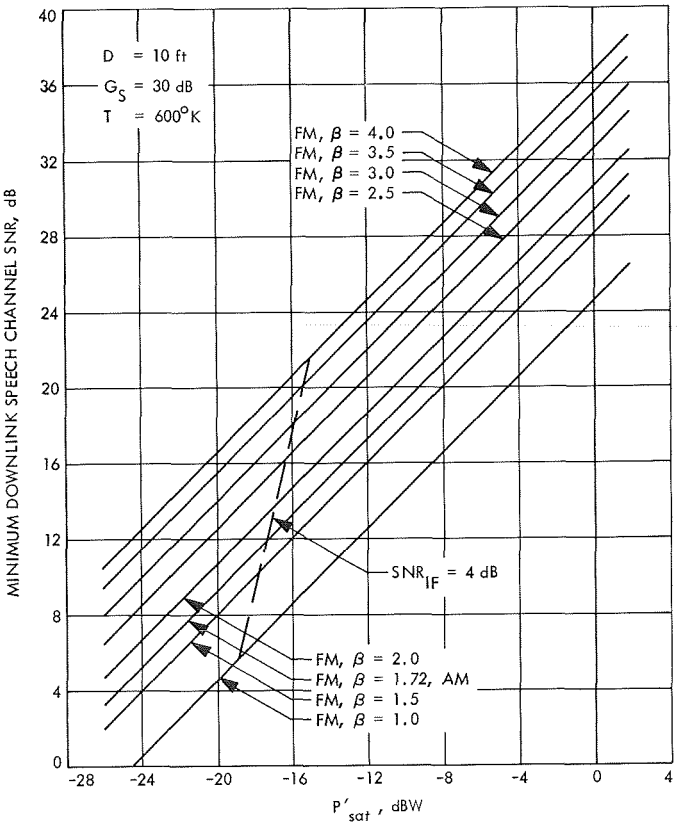


Fig. 10. Downlink SNRs versus satellite power

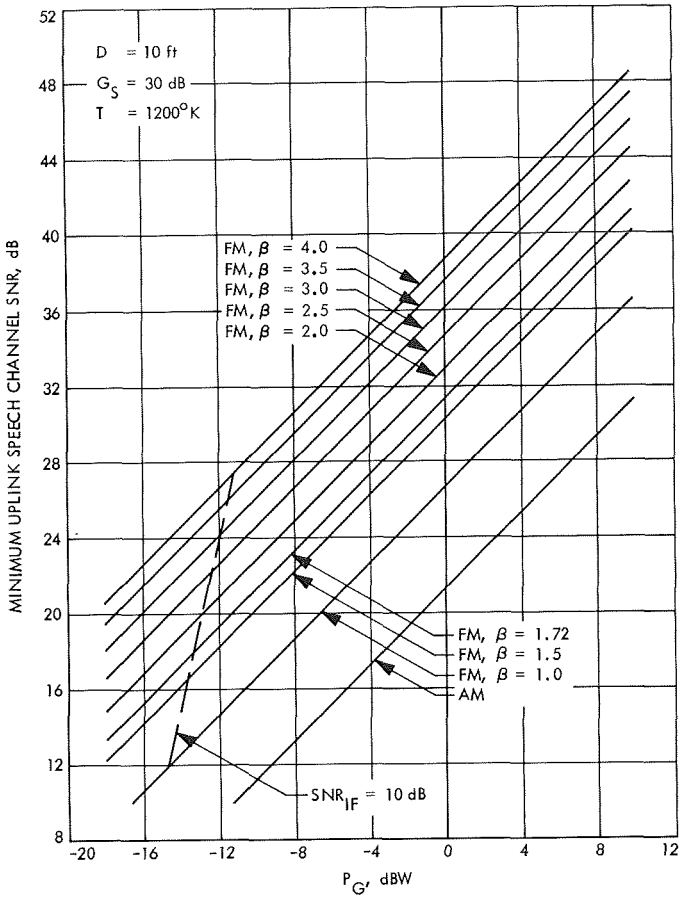


Fig. 11. Uplink SNRs versus ground station power

advantage of FM over AM is 1.6 dB if the design point is 14 dB. The FM advantage is much greater, however, if we choose a high SNR design point. For threshold at 20 dB, $\beta = 3.4$ can be used and the required P'_{sat} is -15.3 dBW. Thus, FM with $\beta = 3.4$ achieves a 6-dB higher speech channel SNR than AM with the same satellite power. Compared to the minimum-power 14-dB system, the minimum-power, 20-dB system requires only 1.5 dB more satellite power. Therefore, it may be worthwhile to consider minimum speech channel SNRs of 20 dB rather than 14 dB, since the cost is only 1.5 dB in satellite power, or, alternatively, in ground station quality. There is, of course, some cost in bandwidth.

We have seen that the downlink advantage of FM over AM is small if the design point is 14 dB, but very significant if the design point is 20 dB. For the uplink, however, FM is distinctly superior at all SNRs. It is obviously desirable to have the uplink SNR large enough so that the uplink causes negligible quality degradation of the two-way signal, even for high-quality receiving stations.

With FM, this is possible with very little transmitter power: $P_G = -6$ dBW (1/4 W) results in a minimum uplink SNR of 27 dB for $\beta = 2.1$ and 31.2 dB for $\beta = 3.4$ for the example system. To get 27 dB with AM would require $P_G = 5.8$ dBW (3.8 W).

b. System bandwidth. The major advantage of AM over FM is in terms of system bandwidth. Single-sideband AM requires only about 4 kHz per channel address, whereas FM requires about 25 kHz with $\beta \approx 2$ and about 35 kHz with $\beta \approx 3.5$. However, even for the widest FM bandwidth, the overall system bandwidth would probably be acceptable since more than 1000 distinct addresses could be assigned in a bandwidth of less than 50 MHz.

c. Uplink power balance. In calculating the downlink transmitter power per active channel, we have assumed that the average uplink received power is the same for each active user. The satellite power is then shared between the active users. (A 2-dB degradation for imbalance has been allowed.) The uplink power balance is easier to accomplish with FM than with suppressed-carrier AM because, with FM, the power is constant whenever a channel is active. For AM, it would be necessary to use automatic gain control (AGC) to keep the short-term average power of each active channel constant. Otherwise, it would be possible for a user to use more than his share of satellite power by overmodulating and allowing his speech to be somewhat clipped by his transmitter with a resulting higher average power. The required AGC might be difficult to mechanize and might cause some quality degradation.

d. Data transmission. Both AM and FM could be used for teletype or other data transmission. Performance would be better with FM because full modulation could be used; with AM, the average uplink power would have to be controlled to prevent a data channel from using more than its share of satellite power. Even then, however, the SNR would be adequate for almost error-free transmission.

7. Overall System Configurations

In this subsection, we discuss overall system configurations and costs for satellites servicing a large number of low-density users.

a. Multiple access methods. Although either FDMA or orthogonal CDMA could be used, FDMA would probably be chosen on the basis of simplicity and cost. If system timing can be provided cheaply, however, CDMA might be cheaper due to the elimination of the need to generate a large number of carrier frequencies. Access could be either random or controlled. For random access, each subscriber would be assigned either a distinct frequency as his address if FDMA is used, or a distinct phase of a PN sequence if CDMA is used. Each speaker would transmit on the address assigned to the listener. This system is basically simple, but has the disadvantages that the ground stations must be able to transmit on a large number of frequencies or PN sequence phases, and that the system bandwidth is wide. For controlled access, channels are assigned on demand. Less bandwidth is used, but operator control is required. All system variations require the same satellite power per active user.

b. Satellite and ground system requirements. As shown in *Subsection 6-a*, an FM system with $\beta \approx 3.4$ requires only 0.03 W (−15.3 dBW) of satellite power per active user to support good quality speech (20 dB) with for ground stations with $D = 10$ ft and $T = 600°$K. With only 6 W of satellite power, 200 active users (100 conversations) could be supported. Even ground stations with $D = 5$ ft could be used, with 24 W required for 100 conversations. A reasonable system could then support ground stations with $D = 5$ ft with good quality and ground stations with $D = 10$ ft with excellent quality.

With AM, ground stations with $D = 5$ ft would attain only fair quality with 24 W of satellite power for 100 conversations, and ground stations with $D = 10$ ft would have good (20 dB) quality.

Using FM for the uplinks, 1 W of transmitter power is sufficient for an uplink SNR of over 31 dB with $\beta = 3.4$, and with only ground stations with $D = 5$ ft. With AM, 10 W would be required even with ground stations with $D = 10$ ft.

c. Satellite availability. By about 1975, space-qualified, solid-state transmitters should be available at frequencies up to about 6 GHz. These transmitters will be about 20–25% efficient, so that RF powers of about 14–15 dBW (25–32 W) should be possible from satellites in the 300-lb class with about 130 W of raw power. These parameters, plus cost per satellite, are similar to the proposed Hughes ETV Satellite (Ref. 2).

Using FM, one 300-lb satellite would provide good quality speech for about 250 active users with ground stations with $D = 5$ ft, or for about 1000 active users with ground stations with $D = 10$ ft. Estimating the peak load to be 25% of the number of subscribers, one satellite could support 1000–4000 users.

Throughout this discussion, we have assumed a 30-dB G_S for a 2000-mi diam earth coverage. Less power is required for less coverage, and more for wider coverage.

d. Ground station costs. Unfortunately, ground station costs are difficult to estimate. Cost estimates are given in Ref. 2 (Table 4-14) for various configurations, but some of these estimates seem quite high for the simple multiple-access method and one-user ground station concept considered here.

8. Conclusions

A satellite system for speech communications between low-density users is definitely technically feasible. The practicality is a matter of economics, with ground-station cost being the primary factor. The success of a system would depend on the number of potential subscribers in a given coverage area to whom a reliable communication system is worth the cost.

References

1. Yeh, L. P., "Multiple-Access Tradeoff Study for Intra-South American Satellite Communication System," *IEEE Trans. Comm. Tech.*, Vol. Com-16, pp. 721–730, Oct. 1968.
2. *Final Report on New Communications Technologies for Less Developed Countries for the President's Task Force on Communications Policy, Volumes 1 and 2*, Page Communications Engineers, Inc., Washington, D.C., Mar. 31, 1969.
3. *Transmission Systems for Communications*, Chapter 10, Bell Telephone Laboratories, 1964.
4. Hirsh, I. J., Reynolds, E. G., and Joseph, M., "Intelligibility of Different Speech Materials," *J. Acoust. Soc. Am.*, Vol. 26, No. 4, pp. 530–538, July 1954.
5. Pickett, J. M., "Effects of Vocal Force on the Intelligibility of Speech," *J. Acoust. Soc. Am.*, Vol. 28, No. 5, pp. 902–905, Sept. 1954.
6. Christman, R. J., Renaud, G. E., and Rubin, A., *The Effect of Speech and Noise Level on the Masking of Speech by Pure Tones and Noise*, Rome Air Development Center Report RADC-TN-60-97 (Defense Documentation Center AD 241096), June 1960.
7. Pollack, I., and Pickett, J. M., "Masking of Speech by Noise at High Sound Levels," *J. Acoust. Soc. Am.*, Vol. 30, No. 2, pp. 127–130, Feb. 1958.
8. CCIR Doc. 11th Plenary Assembly (Green Book), Vol. IV, Part 2, p. 212, Recommendation 358-1, Oslo, Norway, 1966.
9. Schwartz, M., Bennett, W. R., and Stein, S., *Communication Systems and Techniques*, Chapter 3, McGraw-Hill Book Company, Inc., New York, 1966.
10. Fagot, J., and Magne, P., *Frequency Modulation Theory*, Pergamon Press, New York, 1961.

VII. Communications Elements Research

TELECOMMUNICATIONS DIVISION

A. Real-Time Optical Correlation, *A. Owyoung*

1. Introduction

The ever-increasing rates at which data must be transmitted has fostered a need for real-time signal-processing systems capable of handling such rates. Renewed interest in the field of optics spawned by the laser, coupled with the recent development of efficient high-frequency piezoelectric transducers, has introduced the possibility of using opto-acoustic devices to help solve some of these problems of data processing. This report presents a primary feasibility study of the possibilities for using an opto-acoustic system for correlating a received signal against a library of many reference signals and picking out the one which best fits the received signal.

This report is divided into two parts: the possibility of achieving real-time optical correlation is demonstrated in the first part, which explains the system in a heuristic manner; the second part consists of an annotated bibliography which is intended to provide a basis for further work in this area.

In essence, all real-time optical correlators may be visualized by the simple model shown in Fig. 1. A plane light wave incident upon a photographic slide or optical grating with a transmission function $T_1(X_1)$. The slide moves at a velocity v with respect to another slide with a transmission function $T_2(X_2)$.

Since the X_1 frame is moving with respect to the X_2 frame, we may write $X_1 = X_2 - vt$, where t is the elapsed time. The light coming out of the second slide has then been modulated by the two slides and carries the signal

$$T_1(X_1)T_2(X_2) = T_1(X_2 - vt)T_2(X_2)$$

If this signal is now integrated by a light-collecting detector, the output will yield the cross-correlation of the signal on the slides

$$\phi_{T_1T_2}(t) = \int_{-\infty}^{\infty} T_1(X_2 - vt)T_2(X_2)dX_2$$

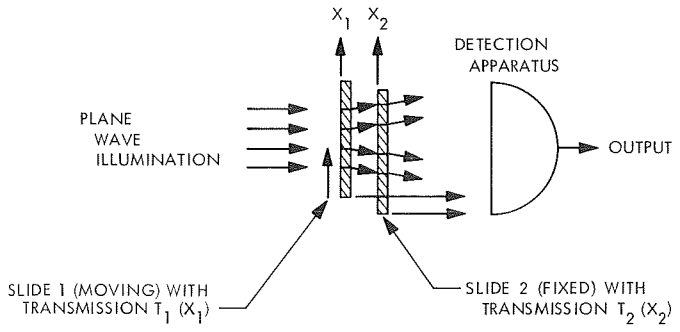


Fig. 1. Real-time optical correlator model

where the domains of $T_1(X_1)$ and $T_2(X_2)$ are finite, so that the limits of integration may be taken at infinity if we assume $T_1 \equiv T_2 \equiv 0$ outside their domains of definition.

This very simple system has a major disadvantage in that the received signal must be recorded on film before processing can be attempted. One means of overcoming this difficulty is to use an ultrasonic modulator in place of the first slide. This would not only overcome the problem of facilitating real-time operation, but would also remove problems associated with having any moving parts in the system. The signal would be introduced as a traveling ultrasonic wave into the crystal column by a piezoelectric transducer and sweep past the reference slide at the velocity of sound in the crystal as shown in Fig. 2.

The system presented in the first part of this report is simply an implementation of the principles just set forth. However, the system is tutorial in nature and is presented only to demonstrate fundamental principles. The references in the second part of this report will thus provide information on how the system might be modified to overcome operational difficulties. It should also be understood that other opto-acoustic correlators exist which are entirely different from the one which is presented in this report. Reference to these are also included.

2. Opto-Acoustic Correlation: Basic Principles of Opto-Acoustic Modulators

Most real-time opto-acoustic correlation systems are based on two fundamental physical principles: (1) the Fourier transform property of a lens, and (2) the phase modulation of light by acoustic waves in material media.

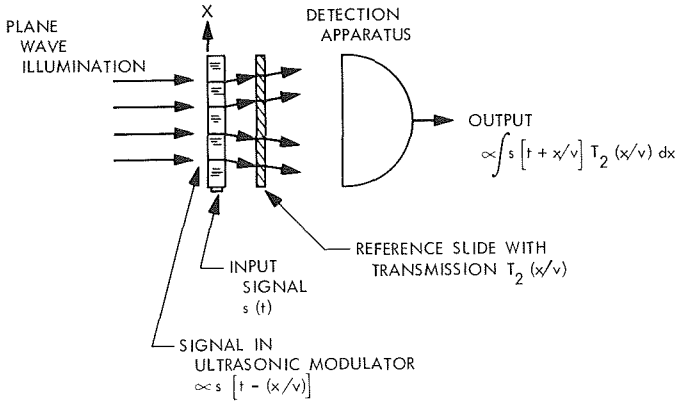


Fig. 2. Ultrasonic light modulator used in real-time optical correlator

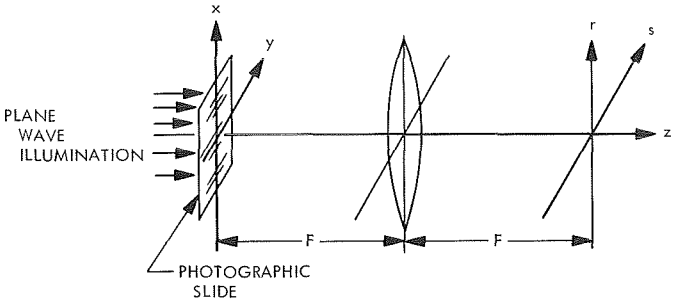


Fig. 3. Two-dimensional spatial Fourier transform relationship between the front and back focal planes of a lens

The existence of a two-dimensional Fourier transform relationship between the front and back focal planes of a convex lens is a well-documented fact (Ref. 1).

This is demonstrated in Fig. 3 where a plane wave of the form e^{ikz} is incident on a photographic slide or optical grating which spatially modulates the wave in the transverse plane (x,y). The Fourier transform property states that the spatial distribution of electric field in the (r,s) plane will be the two-dimensional Fourier transform of the field distribution in the (x,y) plane. This relationship is a rigorous result of the Rayleigh–Sommerfield theory of diffraction, and all of the phase information as well as the amplitude information is retained in the process. More exactly, the relationship is

$$E(r,s,2F) = E(r,s,z)\,\big|_{z=2F}$$

$$= \frac{ike^{ik2F}}{2\pi F}\iint_{-\infty}^{\infty} dx\,dy\,E(x,y,0)\exp\left[-\frac{ik}{F}(rx+sy)\right]$$

$$= \frac{ike^{ik2F}}{2\pi F}E\left(\frac{kr}{F},\frac{ks}{F}\right) \equiv C\,E\left(\frac{kr}{F},\frac{ks}{F}\right) \qquad (1)$$

where

$E(x,y,z)$ = the electric field at the point (x,y,z)

$$C = \frac{ike^{ik2F}}{2\pi F}$$

F = the focal length of the lens

$E(p,q)$ = the Fourier transform of $E(x,y)$.

The limits of integration are taken to be at infinity for convenience. In fact, $E(x,y,0)$ is nonzero only over a finite

aperture and identically zero outside of this aperture; thus the integration need only be taken over the aperture. It is assumed that the lens diameter is much larger than this aperture so that diffraction effects due to the lens may be neglected. As an illustration of the transform relationship, we can clearly see that it holds true in the case where the slide is removed from the (x,y) plane, i.e., the transmission of the slide $T(x,y) = 1$. Ignoring the diffraction effects of the finite lens aperture, we then obtain a plane wave which is focused to a spot at the center of the (r,s) plane. Letting $E(x,y,0) = 1$ in Eq. (1) yields

$$E(r,s,2F) = C\,E\left(\frac{kr}{F}, \frac{ks}{F}\right)$$

$$= C\,\delta\left(\frac{kr}{F}\right)\delta\left(\frac{ks}{F}\right) = C'\,\delta(r)\,\delta(s) \qquad (2)$$

where

$\delta(r)$ = the dirac delta function

$$C' = C\frac{F^2}{k^2}$$

The modulation of a plane wave, e^{ikz}, by an ultrasonic wave in a crystal can be seen by recalling that the wave vector **k** has an amplitude which is linearly dependent on n, the index of refraction of the medium.

$$|k| = \frac{\omega}{c}n \qquad (3)$$

An ultrasonic wave propagating perpendicular to the optic axis, as shown in Fig. 4, will cause a variation of the index of refraction of the crystal due to the compression

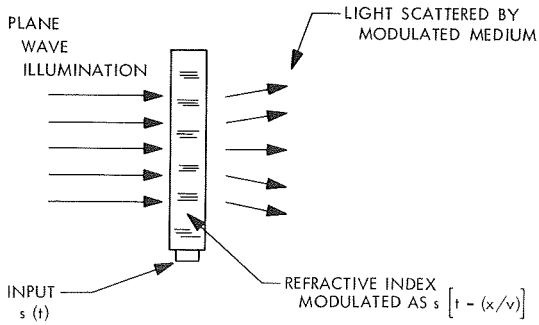


Fig. 4. Piezoelectric transducer deposited on a crystal to produce an ultrasonic light modulator

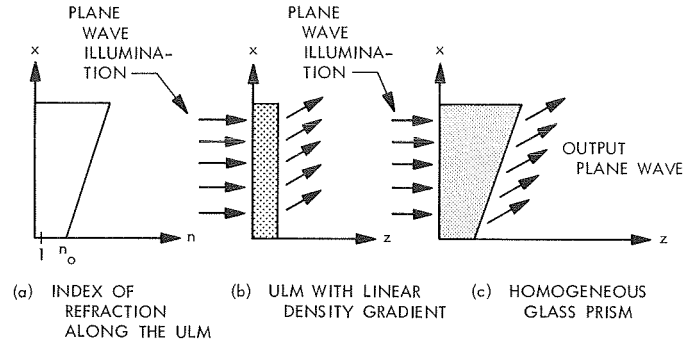


Fig. 5. Bending of incident light rays by spatial gradients: (a) Index of refraction along the ULM, (b) linear density gradient of ULM, (c) homogeneous glass prism delaying portions of the wavefront

and rarefaction of the material. An acoustic input signal $s(t)$ will thus propagate in the crystal as a modulation of the index

$$n = n_0 + \alpha s\left(t - \frac{x}{v}\right)$$

where

v = the velocity of sound in the medium

α = a scaling factor

n_0 = the average index

The resultant modulation of the plane wave will thus produce fields at the output of the ultrasonic light modulator (ULM) which are of the form

$$E(x,y,0) = E_0' \exp\left[i\frac{\alpha}{n_0}k_0 s\left(t - \frac{x}{v}\right)\right] \qquad (4)$$

where

$$k_0 = \frac{\omega n_0}{c}$$

E_0' = a lumped constant

We can thus envision the modulated index to be spatial gradients in the optical path length which delay certain portions of the wavefront longer than others. These gradients may also be viewed as a means by which the incident light rays are bent or scattered as they traverse

the medium. This is demonstrated in Fig. 5 for a ramp function modulation of the index.

We see that the delay produced by the increase in index in the x-direction is equivalent to lengthening the path. It is also interesting to observe that the upward bending of the light rays in Fig. 5b corresponds to an upward spatial frequency shift; for, if we had replaced the slide in Fig. 3 with the ULM, the upward wave out of the ULM would produce a focused spot above the optic axis in the focal plane of the lens. The ULM is thus seen to have truly "modulated" some energy into an upper sideband.

Applying this same type of argument to the general case in Eq. (4) and assuming that $s(t)$ is a narrowband signal centered about its carrier frequency Ω_0, we recognize Eq. (4) to be a phase modulation of the plane wave. Its Fourier spectrum will thus show many sidebands at the focal plane of the lens, and many spots will result in this plane. We note that if $s(t)$ were a purely sinusoidal signal, the result in the transform plane would be a number of spots the amplitude of the nth order spot being proportional to the nth order Bessel function J_n. It is also to be observed that since the signal $s(t)$ travels up the ULM at velocity v, the output in the transform plane is also time varying.

Now that the spectrum of the signal is available in the transform plane, we are able to filter the signal or modify it in various ways by blocking or attenuating various diffraction orders of the scattered light. A reimaging of the ULM output after the filtering process will then display the image with the blocked-out frequencies removed. We will see that this method of "spatial filtering" will prove to be a valuable tool.

a. A possible opto-acoustic correlator. One of the most basic opto-acoustic correlators (Fig. 6) is that proposed in Ref. 2.

In this system a plane wave illuminates the ULM which is modulated by a signal $s(t) = a(t)\ cos(\Omega_0 t)$, where $a(t)$ is a baseband signal with bandwidth Ω such that $2\Omega < \Omega_0$. If the modulation is weak, i.e., if α is sufficiently small, Eq. (4) may be expanded in powers of α and we may write:

$$
\begin{aligned}
E(x,y,0) &= E_0' \exp\left(i\frac{\alpha}{n_0} k_0 a\left[t - \frac{x}{v}\right] \cos\left[\Omega_0\left(\frac{x}{v} - t\right)\right]\right) \\
&\approx E_0'\left[1 + \frac{\alpha k_0}{n_0} a\left(t - \frac{x}{v}\right) \cos\left(Kx - \Omega_0 t\right)\right] \\
&= E_0'\left[1 + \frac{\alpha k_0}{2n_0} a\left(t - \frac{x}{v}\right)\right. \\
&\quad \left. \times \left(\exp[i(Kx - \Omega_0 t)] + \exp[-i(Kx - \Omega_0 t)]\right)\right]
\end{aligned}
\tag{5}
$$

where

$$K = \frac{2\pi}{\Lambda}$$

Λ = acoustic wavelength

From Eq. (5) it is clear that there is a component of the plane wave which passes through the ULM unaffected and that some energy is scattered into the first upper and lower spatial sidebands which carry the signal $a(t)$. Lens 1 then Fourier transforms the signal. The fundamental beam along with the lower order sidebands are blocked out by the knife edge in the focal plane of this lens. The

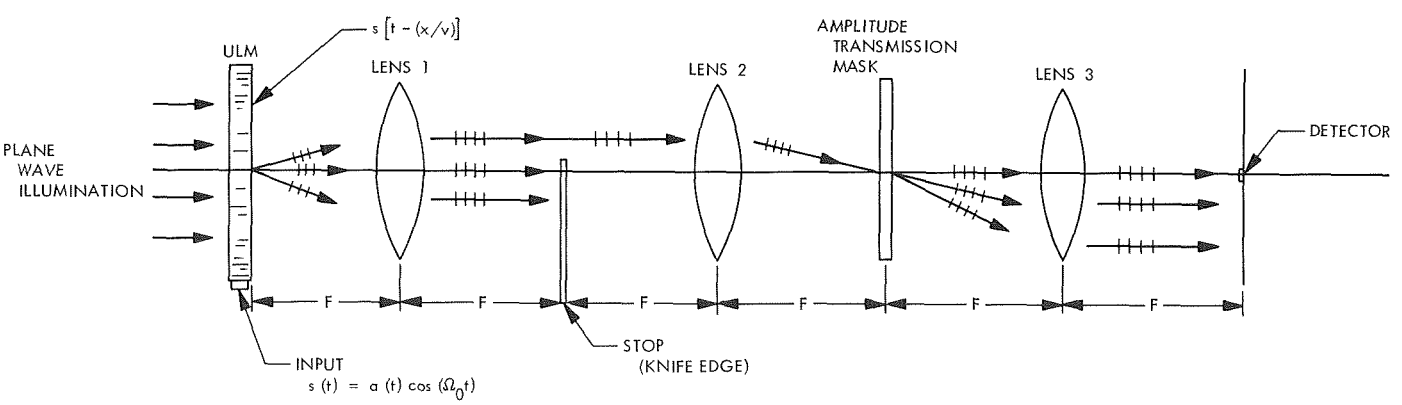


Fig. 6. A possible opto-acoustic modulator. Arrows indicate directions of fundamental and first-order diffracted beams

signal centered about the upper sideband is allowed to pass on through the system, and lens 2 forms an image of the upper sideband signal

$$a\left(t-\frac{x}{v}\right)\exp[i(Kx-\Omega_0 t)]$$

on the amplitude mask. The knife edge has thus performed the function of transforming our phase-modulated signal into an amplitude-modulated signal. If the amplitude mask has an amplitude transmission of the form

$$T(x)=T_0+b\left(\frac{x}{v}\right)\cos(Kx)$$

where T_0 is a bias term (since T cannot be negative), the output will be of the form

$$E(x,y,4F)\propto a\left(t+\frac{x-\ell}{v}\right)\exp[-i(Kx+\Omega_0 t)]\times\left[\frac{b}{2}\left(\frac{x}{v}\right)(e^{iKx}+e^{-iKx})+T_0\right] \quad (6)$$

where ℓ is the length of the ULM along the acoustic axis and the argument of the first factor has been altered by the lenses. It should be clear that this result is again a modulation of an input wave which generates sidebands. The only differences are that the amplitude modulation results in only two sidebands and the input wave is now coming in at a slight angle, since the input signal is being carried on an upper sideband. Lens 3 now takes the Fourier spectrum of the mask output. The Fourier spectrum of Eq. (6) shows that one of the sidebands generated by the mask produces the autocorrelation at $x,y=0$ in the output plane. More specifically

$$E(x,y,5F)\propto\int_{-\infty}^{\infty}a\left(t+\frac{x-\ell}{v}\right)b\left(\frac{x}{v}\right)\exp\left(\frac{-ikxx'}{F}\right)dx\Bigg|_{x'=0}\triangleq\phi_{ab}\left(t-\frac{\ell}{v}\right) \quad (7)$$

where $\phi_{ab}(\tau)$ = the cross-correlation between a and b. Calculation will show that other terms will only make contributions to other points on the x,y plane. This is also intuitively reasonable from Fig. 6. Lens 3 thus performs the integration, with optical detection being performed at $x,y=0$ in the output plane. In essence, lens 3 integrates the fields as the image of the ULM is "swept" past the amplitude mask.

It is worthwhile to make one final observation. The output current will be the square of the correlation function, due to the square-law nature of optical detectors. It can be shown, however, that a linear output may be achieved through heterodyning techniques whereby another plane wave is introduced at the output plane which is phase-locked to the input wave. The beating of this wave with the output wave will give a linear output (Ref. 3)

3. Annotated Bibliography

a. Acoustic diffraction. The diffraction of light by ultrasonic waves in material media was first predicted by Brillouin in 1921 (Ref. 4) and verified independently by the experimental work of Sears and Debye (Ref. 5) and Lucas and Biquard (Ref. 6). The simultaneous occurrence of many orders in the diffraction process was first explained by Raman and Nath in a series of comprehensive articles on the interaction of light and sound (Ref. 7). The frequency shift of the light predicted by the Raman-Nath theory was experimentally verified by Cummins et al. (Ref. 8). For a tutorial review of the physical concepts underlying ultrasonic light diffraction the reader is referred to the article by Adler (Ref. 9) and the book by Born and Wolf (Ref. 10).

In this report we have treated only the case of the phase modulation of light by sound. This is referred to as ultrasonic diffraction of light in the "Raman-Nath regime" or "Sears-Debye regime." Under somewhat different conditions, when the width of the acoustic wave is large, the frequency high, or the amplitude high, amplitude modulation of the light may also result. We then consider diffraction in the Bragg regime. The criteria for determining the regime dictated by a given set of experimental conditions is often a subject of confusion. This topic is treated by Willard (Ref. 11). The papers by Klein and Cook (Refs. 12 and 13) also give an excellent account of the problem in addition to considering the transition region between the two regimes. Probably the most exhaustive and generalized treatment of ultrasonic light diffraction is given by Bhatia and Noble (Ref. 14). This approach is summarized in Ref. (10).

Unlike Raman–Nath diffraction, Bragg diffraction is not as amenable to straightforward mathematical analysis. Hargrove treats this problem as a series of N successive Raman-Nath type of diffractions (Ref. 15) and Klein and Hiedeman (Ref. 16) have performed an experimental investigation of the theory developed by Mertens (Ref. 17). A review of the subject is given by Quate et al. (Ref. 18)

along with a rather extensive bibliography. Dixon (Ref. 19) extends the theory to the case of anisotropic crystals.

b. Ultrasonic light modulation. The field of opto-acoustic modulators is a rapidly growing one, and many breakthroughs are being made in the area of piezoelectric transducers. Rather than making an artificial attempt at citing the mass of publications concerned with the various materials and techniques which have been developed over the years, a few basic papers with comprehensive bibliographical material will be cited.

Early work concerned with the details of photoelastic interactions in crystals was performed by Mueller (Refs. 20 and 21). The papers by Maloney et al. might serve as a starting point for material concerning the nature of light and sound diffraction in crystals (Refs. 22 and 23).

The literature concerned directly with ULM's came somewhat later. In 1947, Arenberg published an excellent paper dealing with both theoretical and experimental aspects of ultrasonic solid delay lines (Ref. 24). Gordon considers Bragg modulators in his review of the ULM (Ref. 25) and performs sample calculations on the device used as an amplitude modulator. More recently, McMahon considered the efficiency of the Bragg modulator and its dependence on geometry (Ref. 26). The response characteristics of the Raman-Nath modulator is considered in an excellent paper by Minkoff, et al. (Ref. 27).

Materials research in acoustics is quite active, and a review of some of the newer opto-acoustic materials is found in a paper by Spencer et al. (Ref. 28). Piezoelectric transducer materials are also discussed in a review by Jaffe et al. (Ref. 29). The entire journal in which Ref. 29 is found is devoted to the subject of ultrasonics.

c. Correlation systems. Let us now consider real-time signal processors, of which optical correlation processors are one specific application. For a review of other systems which employ photographic film to record the signal prior to processing, the reader is referred to Ref. 1 and to the excellent review paper by Vander Lugt (Ref. 30).

Most of the early work on real-time optical correlation employs some spatial filtering technique (Ref. 31). One of the first such systems was proposed by Slobodin (Ref. 32) and analyzed by Felstead, who also gives some interesting experimental data on the system (Ref. 33). A similar, but somewhat more sophisticated, technique was proposed by Arm et al. (Ref. 34). Felstead also discusses this system (Ref. 35). The above correlators employ a fixed reference function to be correlated against a variable input. Izzo (Ref. 36) proposes a system with a variable reference. Palfreeman also proposes a system with a variable reference in a rather interesting article which relies heavily on intuition arguments (Ref. 37). In place of a spatial filtering technique such as the one used in the first part of this paper, a rather novel method of polarization discrimination is used to separate the fundamental beam from the "signal carrying" sideband beam.

More recently, Meltz and Maloney have developed a correlator which employs the principle of Fresnel images. An extensive analysis of their system is given in Ref. 38. Carleton, Maloney, and Meltz have compared this system to some of the systems cited above in their highly enlightening article on heterodyne correlators (Ref. 3).

Thus far, all the systems proposed have employed Raman-Nath modulators. Bragg modulators usually operate at a higher frequency and thus would be expected to have larger bandwidths. Jernigan has proposed a system which employs a Bragg modulator and operates with a fixed reference function (Ref. 39). Squire and Alsup have proposed a similar system, which use two ULM's and thus employs a variable reference function (Ref. 40).

d. Summary and conclusion. In surveying the literature on real-time optical correlators, it was found that the area was relatively new and still in its initial developmental stages. Most experiments were concentrated on obtaining an operational system rather than an optimal one and very little information is available on system capabilities. The major system limitations arise from the capability of the transducers. Palfreeman (Ref. 37) predicts operation at 60–80 MHz with 50% bandwidth and a 200-μs signal length. Squire (Ref. 40) predicts a 500-MHz bandwidth with a time bandwidth product of 2000 for his system employing Bragg modulators. Further experimental work is being done to realize these limits. With the advances being made in acoustic technology, it seems worthwhile to further investigate this area.

Since the development of opto-acoustic devices correlation is a rather recent one, there is still quite a bit of confusion in the literature, e.g., Ref. 33. Any further work should thus start with a thorough analysis of the prospective systems to fully understand their operation, capabilities, and limitations.

References

1. Goodman, T. W., *Introduction to Fourier Optics,* McGraw-Hill, Inc., New York, 1968.
2. King, M., et al., "Real Time Electrooptical Signal Processor With Coherent Detection," *App. Optics,* Vol. 6, pp. 1367–1375, Aug. 1967.
3. Carleton, H. R., Maloney, W. T., and Meltz, G., "Colinear Heterodyning in Optical Processors," *Proc. IEEE,* Vol. 57, pp. 769–775, May 1969.
4. Brillouin, L., "Diffusion de la lumiere et des rayons X par un corps transparent homogene," *Ann. Phys.* (Paris), 9th Ser., Vol. 17, p. 88, 1922.
5. Debye, P., and Sears, F. W., "On the Scattering of Light by Supersonic Waves," *Proc. Nat. Acad. Sci.,* (USA), Vol. 18, p. 409, 1932.
6. Lucas, R., and Biquard, P., "Optical Properties of Solids and Liquids Under Ultrasonic Vibrations," *J. Phys. Rad.,* 7th Ser., Vol. 3, p. 464, 1932.
7. Raman, C. V., and Nath, N. S. N., "The Diffraction of Light by High Frequency Sound Waves: Part I," *Proc. Indian Acad. Sci.,* Vol. 2, pp. 406–412, 1935; "Part II," Vol. 2, pp. 413–420, 1935; "Part III," Vol. 3, pp. 75–84, 1936; "Part IV," Vol. 3, pp. 119–125, 1936; "Part V," Vol. 3, pp. 459–465, 1936; "Generalized Theory," Vol. 4, pp. 222–242, 1936.
8. Cummins, H., et al., "Frequency Shifts in Light Diffracted by Ultrasonic Waves in Liquid Media," *App. Phys. Lett.,* Vol. 2, pp. 62–64, Feb. 1963. See also Cummins, H., and Knable, N., "Single Sideband Modulation of Coherent Light by Bragg Reflection From Acoustic Waves," *Proc. of IEEE,* (correspondence), Vol. 51, p. 1246, Sept. 1963.
9. Adler, R., "Interaction Between Light and Sound," *IEEE Spectrum,* pp. 42–54, May 1967.
10. Born, M., and Wolf, E., *Principles of Optics,* Ch. 12, Sect. 12.1, Pergamon Press, New York, 1965.
11. Willard, G. W., "Criteria for Normal and Abnormal Ultrasonic Light Diffraction Effects," *J. Acous. Soc. Am.,* Vol. 21, No. 2, pp. 101–108, February 1949.
12. Klein, W. R., Cook, B. D., and Mayer, W. G., "Light Diffraction by Ultrasonic Gratings," *Acoustics,* Vol. 15, No. 2, pp. 67–74, Jan. 1965.
13. Klein, W. R., and Cook, B. D., "Unified Approach to Ultrasonic Light Diffraction,: *IEEE Trans. Sonics Ultrasonics,* Vol. SU-14, pp. 123–134, July 1967.
14. Bhatia, A. B., and Noble, W. J., "Diffraction of Light by Ultrasonic Waves, I. General Theory, II. Approximate Expressions for the Intensities and Comparison With Experiment," *Proc. Royal Soc.,* Vol. 220, pp. 356–385, May 1953.
15. Hargrove, L. E., "Optical Effects of Ultrasonic Waves Producing Phase and Amplitude Modulation," *J. Acous. Soc.,* Vol. 34, pp. 1547–1552, Oct. 1962.
16. Klein, W. R., and Hiedeman, E. A., "An Investigation of Light Diffracted by Wide High Frequency Ultrasonic Beams," *Physica,* Vol. 29, pp. 981–986, 1963.
17. Mertens, R., *Meded. Koninklijke Vlaamse Academie Wetenschappen* Vol. 12, No. 7, 1950.
18. Quate, C. F., Wilkinson, C. D. W., and Winslow, D. K., "Interaction of Light and Microwave Sound," *Proc. IEEE,* Vol. 53, pp. 1604–1623, Oct. 1965.
19. Dixon, R. W., "Acoustic Diffraction of Light in Anisotropic Media," *IEEE J. of Quant. Elec.,* Vol. QE 3, pp. 85–93, Feb. 1967.
20. Mueller, H., "The Intensity and Polarization of the Light Diffracted by Supersonic Waves in Solids," *Phys. Rev.,* Vol. 52, pp. 223–229, Aug. 1, 1937.
21. Mueller, H., "Determination of Elastooptical Constants With Supersonic Waves," *Z. Kryst,* Series A, Vol. 99, pp. 122–141, 1938.
22. Maloney, W. T., Carleton, H. R., "Light Diffraction by Transverse Ultrasonic Waves in Hexagonal Crystals," *IEEE Trans. Sonics Ultrasonics,* Vol. SU-14, pp. 135–139, July 1967.
23. Maloney, W. T., Meltz, G., and Gravel, R. L., "Optical Probing of the Fresnel and Fraunhofer Regions of a Rectangular Acoustic Transducer," *IEEE Trans. Sonics Ultrasonics,* Vol. SU-15, pp. 167–172, July 1968.
24. Arenberg, D. L., "Ultrasonic Solid Delay Lines," *J. Acoustic Soc. Am.,* Vol. 20, pp. 1–25, Jan. 1948.
25. Gordon, E. I., "A Review of Acoustooptical Deflection and Modulation Devices," *Proc. IEEE,* Vol. 54, pp. 1391–1401, October 1966.
26. McMahon, D. H., "Relative Efficiency of Optical Bragg Diffraction as a Function of Interaction Geometry," *IEEE Trans. Sonics Ultrasonics,* Vol. SU-16, pp. 41–44, April 1969.
27. Minkoff, J. B., et al., "Transfer Characteristics and Frequency Response of Wideband Amorphous-Solid Debye-Sears Light Modulator," *J. App. Phys.,* Vol. 39, pp. 2024–2029, March 1968.
28. Spencer, E. G., Lenzo, P. V., and Ballman, A. A., "Dielectric Materials for Electrooptic, Elastooptic and Ultrasonic Device Applications," *Proc. IEEE,* Vol. 55, pp. 2074–2108, Dec. 1967.
29. Jaffe, H., and Berlincourt, D. A., "Piezoelectric Transducer Materials," *Proc. IEEE,* Vol. 53, pp. 1373–1386, Oct. 1965.
30. Vander Lugt, A., "A Review of Optical Data-Processing Techniques," *Optics Acts,* Vol. 15, pp. 1–33, 1968.
31. Lambert, L. B., in *Modern Radar.* Edited by R. S. Berkowitz, John Wiley & Sons, New York, 1965.
32. Slobodin, L., "Optical Correlation Technique," *Proc. IEEE,* Vol. 51, p. 1782, Dec. 1963.
33. Felstead, E. B., "A Simple Real Time Incoherent Optical Correlator," *IEEE Trans. Aero. Electron. Sys.,* Vol. AES-3, pp. 907–914, Nov. 1967.
34. Arm, M., Lambert, L., and Weissman, I., "Optical Correlation Technique for Radar Pulse Compression," *Proc. IEEE,* Vol. 52, p. 842, July 1964.
35. Felstead, E. B., "A Simple Coherent Optical Correlator," *Applied Optics,* Vol. 7, pp. 105–108, Jan. 1968.
36. Izzo, N. F., "Optical Correlation Technique Using a Variable Reference Function," *Proc. IEEE,* Vol. 53, pp. 1740–1741, Nov. 1965.
37. Palfreeman, J. S., "An Opto-acoustic Cross Correlator in Radar Signal Detection," *Philips Tech. Rev.,* Vol. 28, pp. 217–225, 1967.

38. Meltz, G., and Maloney, W. T., "Optical Correlation of Fresnel Images," *App. Optics*, Vol. 7, pp. 2091–2099, Oct. 1968.

39. Jernigan, J. L., "Correlation Technique Using Microwaves," *Proc. IEEE*, Vol. 56, p. 374, March 1968.

40. Squire, W. D., and Alsup, J. M., "A Wideband Optical Correlator and Matched Filter Using Diffraction of Light by Ultrasonic Waves," Report NUWC-TP-71, Naval Undersea Warfare Center, Pasadena, Calif. May 1968.

B. Spacecraft Antenna Research: Large Spacecraft Antennas (Non-Paraboloidal Reflector),

A. Ludwig

1. Introduction

Large spacecraft antennas must necessarily be erectable because of limited launch vehicle shroud sizes (Table 1). In addition, a folded structure will generally withstand launch stresses better than a fully extended structure. A major goal in the design of an erectable reflector antenna is to ensure that the surface will conform closely to the desired shape, which is typically determined by RF considerations.

The basic purpose of a reflector is to collect energy over a large region of space, and focus it toward a small region where the antenna feed is located (a feed will be defined as any device that couples energy from free space to a transmission line). For spacecraft, the incident energy is in the form of a plane wave. The natural choice for the surface shape is a paraboloid, which optically converts a plane wave into a spherical wave converging to a point. This is very desirable from an RF standpoint, because it is easy to build a feed which couples well with a converging spherical wave. It is not so desirable from a structural standpoint, because the paraboloid is a doubly curved surface.

A singly curved surface has the inherent disadvantage of focusing in one dimension only, and the best one can hope to do is to convert a plane wave into a cylindrical (or conical) wave which converges toward a line segment.

Table 1. Maximum package diameter for typical launch vehicle envelopes

Booster	Maximum package diameter, ft
Agena	4.8
Atlas or Titan	9.0
Saturn V	20.0

(A paraboloid is, of course, the unique surface which focuses to a point, and some doubly-curved surfaces—such as a segment of a sphere—also focus on a line segment.) This means that one needs either a feed that couples to a cylindrical wave—a "line source" feed—or an additional reflecting surface to re-focus the energy.

It is the intent of this article to present some designs in which the main energy-collecting surface is singly curved. The feed design problem will not be considered here, except to note that it is more difficult to obtain a good line-source feed than a good point-source feed. A second restriction in this preliminary study is that all reflector systems will be based on optical design principles; diffraction considerations will be deferred to a future study.

2. Cylindrical Reflectors

The unique cylindrical reflector which focuses on a line is a parabolic cylinder; this is a common antenna reflector and may be used in conjunction with a line-source feed.

An interesting variation on this standard design is the cylindrical parabolic cassegrain antenna illustrated in Fig. 7. This system focuses an incident plane wave to a point, like a conventional circular parabolic cassegrain system. The main reflector is described by the equation

$$r(\theta,z) = r(\theta) = f \sec^2 \frac{\theta}{2} \tag{1}$$

The subreflector is described by

$$r(\theta,z) = \frac{z^2 + 4(c^2 - a^2)}{4(c\cos\theta + a)} \tag{2}$$

For z = constant, the contour of the subreflector is hyperbolic, with the conventional constants given by

$$c'(z) = \alpha c$$

$$a'(z) = \alpha a$$

where

$$\alpha(z) = 1 + \frac{z^2}{4(c^2 - a^2)} \tag{3}$$

For θ = constant, the contour is parabolic, with focal length

$$f(\theta) = c\cos\theta + a \tag{4}$$

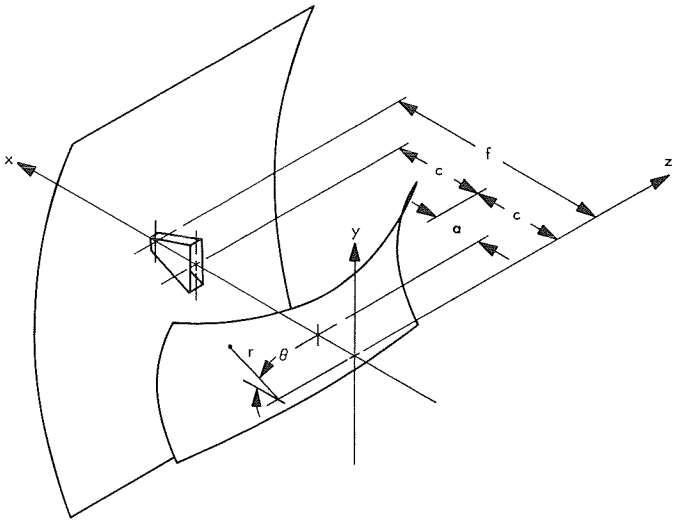


Fig. 7. Cylindrical Cassegrain antenna

It should be noted that the reflectors have been drawn with rectangular boundaries for illustrative purposes only, and, in practice, they would probably be truncated at an elliptical boundary. A second variation on this design is the equivalent Gregorian system in which the subreflector curvature in the Y-direction is reversed.

3. Conical Reflectors

Any cone will focus a plane wave to a line in the form of a conical wave; for a cone included angle of 90 deg the wave becomes cylindrical. Again, this reflector may be used by itself by employing a line source feed. Alternatively, a subreflector may be used to re-focus the energy to a point.

One such reflector/subreflector system, in which the "subreflector" actually is an integral part of the main reflector, is shown in Fig. 8. A more conventional system is shown in Fig. 9. This later case is illustrated in the Gregorian form, and again there is a similar cassegrain form. In both of these designs, all surfaces are figures of revolution about the z-axis, and in both cases the contour of the secondary reflector is parabolic. Another feature that both designs have in common is that the main reflector aperture illumination has a ring-like pattern; i. e., there is a hole in the center. This hole extends to slightly more than half the radius of the main reflector, and the illuminated portion has about 72% of the area of the full aperture. This means that this type of antenna has a *maximum* gain that is 72% of the *maximum* gain of a full circular aperture of the same diameter. However, in a conventional cassegrain or Gregorian system the subreflector blocks an area of perhaps 1–2% of the total area; in addition, this area is illuminated by roughly 2–4% of the total feed power; so the net blockage loss in a conventional system is 3–6%. Therefore, the relative potential gain of the designs given here is about 75–78% of the standard value. Furthermore, the design shown in Fig. 9 has the interesting property of reflecting the center of the feed pattern to the *outer edge* of the main reflector; this will almost certainly result in improved illumination efficiency (this would have to be checked using analysis beyond the scope of this preliminary study), which would narrow the difference further. In any case, the basic motivation of these designs is structural rather than electrical, and any comparisons obviously must include structural effects.

4. Summary and Future Work

Several novel reflector concepts have been presented, with the common feature that the main reflector is a singly curved surface. These designs will be reviewed from a structural standpoint to determine how the surface accuracy and weight compare to conventional systems. Any design that appears structurally promising will then

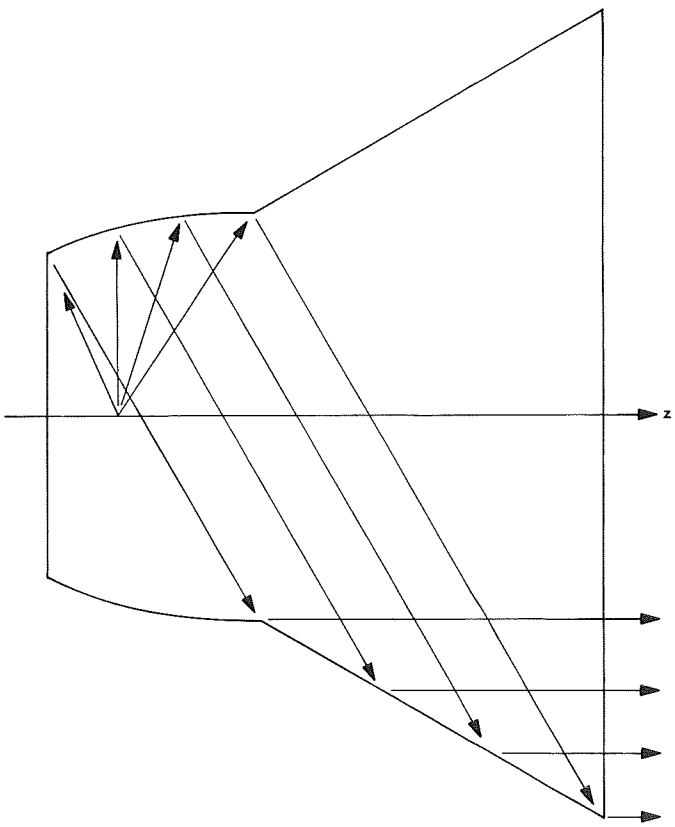


Fig. 8. Conical reflector with integral parabolic subreflector

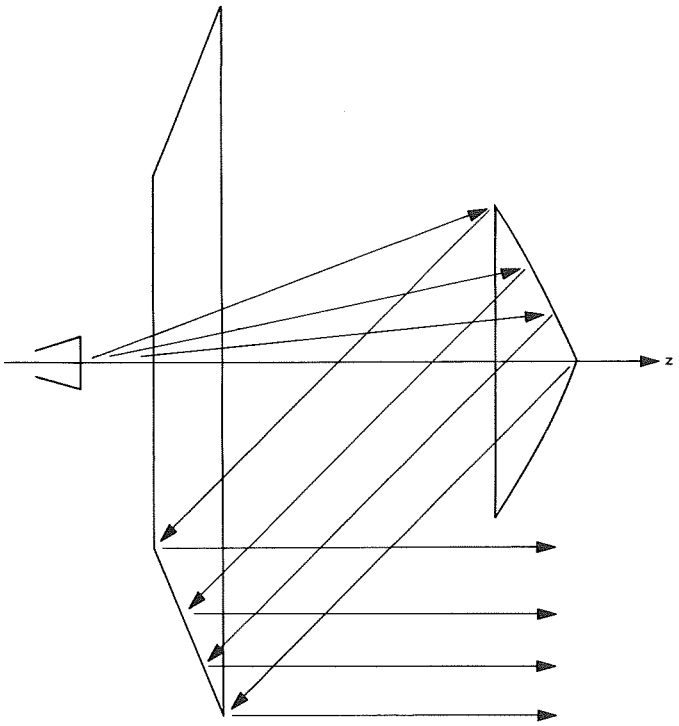


Fig. 9. Conical reflector with Gregorian-type parabolic subreflector

be carefully analyzed from an RF standpoint, including diffraction effects and feasibility of feed design, to determine if any of these designs are an improvement over conventional paraboloidal reflector systems.

C. Spacecraft Antenna Research: Large Spacecraft Antennas (Zone-Plate Lens Antenna Study), *D. MacQuigg*

1. Introduction

The need for higher gain in deep space communication links and the size limitations imposed on the spacecraft by the launch vehicle make the use of erectable antennas a necessity. (See Table 1). A critical problem in the design of erectable antennas is the accuracy of positioning of the antenna elements. In the case of a large reflector, the ray path error is twice the surface error. Use of a lens instead of a reflector would seem an attractive alternative, since the ray path length through a lens system is not as sensitive to surface errors.

Lenses have been used successfully at microwave frequencies (Ref. 1), but a conventional lens the size of a spacecraft antenna would require large volumes of dielectrically dense material and would be too cumbersome to use as an erectable structure. The parallel plate lens would offer some savings in weight at a cost of greater complexity.

Unlike the dielectric and parallel plate lenses which focus radiation by continuously modifying the phase of a wavefront, the zone plate, in effect, synthesizes the desired wavefront by rejecting radiation from zones which are out of phase with the desired wavefront. This rejection of out-of-phase components and the fact that the phase is not uniform over the in-phase zones results in an inherently low efficiency. The zone plate also differs from reflectors and true lenses in that its gain is limited to a narrow frequency band.

2. Principle of the Zone Plate

Consider a circular aperture in an opaque screen, uniformly illuminated with a plane wave on the aperture axis. The diffracted field at a point on axis behind the aperture is a function of aperture radius and wavelength. For a given wavelength, the field, plotted in the complex plane for apertures of increasing radius, is a converging spiral, as shown in Fig. 10a. Successively larger increments in radius result in the addition of small components of increasing phase and diminishing magnitude. The resultant field has been normalized in magnitude and phase to the incident field. As expected, the spiral converges to (1,0) for very large apertures. The case shown here is for the wavelength equal to one fifth the distance

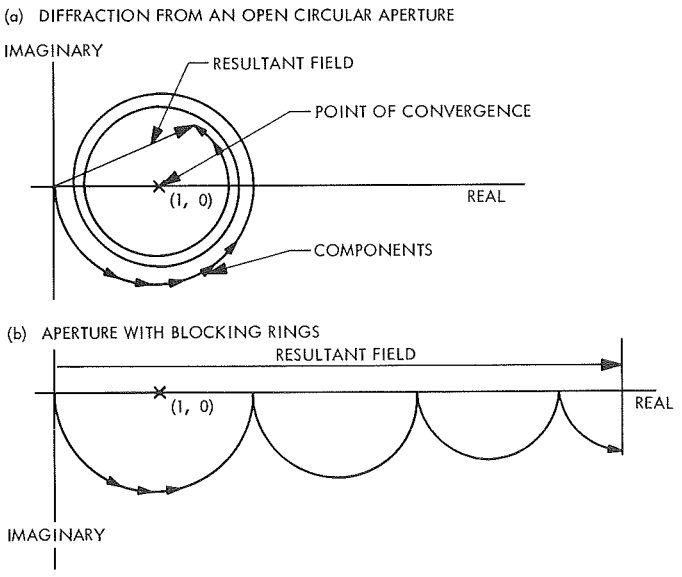


Fig. 10. Diffracted field: (a) from an open circular aperture, (b) from an aperture with blocking rings

from aperture to field point. Smaller wavelength results in a more slowly converging spiral.

The aperture can be divided into Fresnel zones, which are annular regions, concentric about the aperture axis, across which the phase changes by 180 deg, as seen from the field point. The resultant field can be maximized by blocking out radiation from alternate zones. The fields then add as shown in Fig. 10b. The incident wave has, in effect, been focused, even though no phase modification has occurred, as in a true lens. Notice that the resultant field for each zone is $(2/\pi)$ times the total summed magnitude of all the components from that zone.

3. Efficiency of an Ideal Zone Plate

A useful figure of merit for an antenna with a well-defined aperture is the efficiency, defined as the ratio of the power intensity at the focal point (or far-field point) to the intensity that would result if the field across the aperture were uniform in amplitude and phase, with respect to the focal point (or far-field point). An ideal zone plate will be small enough that variations in distance to the focal point will not result in significant amplitude taper, yet large enough that the number of zones is sufficient to ensure that exactly 50% of the aperture area is blocked. The amplitude at the focus will then be reduced 50% due to blocked aperture area and another factor of $(2/\pi)$ due to nonuniform phase across the open zones. The total reduction in amplitude is then $(1/\pi)$, and the efficiency of an ideal zone plate is

$$\left(\frac{1}{\pi}\right)^2 = 10.1\%$$

4. Application to Spacecraft Antenna

In spite of low efficiency of the zone plate, it may be that the simplicity of construction and large error tolerances will allow such an increase in size as to more than make up the loss of efficiency. Assuming the same feed efficiency, a zone plate would be one tenth as efficient as a paraboloid reflector, and would therefore have to be at least three times the diameter to achieve the same gain.

In principle, it should be possible to construct a zone plate on any surface. For example, the antenna might be a large balloon with conducting rings painted on the surface. Buskirk and Hendrix tested a planar structure and achieved 2 to 3% efficiencies using a low-efficiency feed (Ref. 2). One purpose of this study is to determine if the zone plate will function on a spherical surface and to account for the various factors in the loss of efficiency in this case.

5. Construction of the Antenna

The test antenna is shown in Figs. 11 and 12. The feed is a short section of waveguide with an open end 1.00 by 0.50 in., operating at a wavelength of 1.4 in. (8448 MHz), which provides a broad pattern of nearly uniform phase and amplitude.

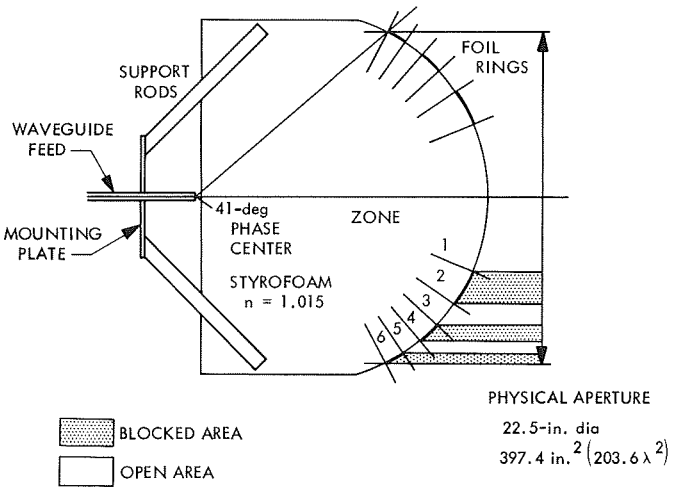


Fig. 11. Zone-plate antenna schematic

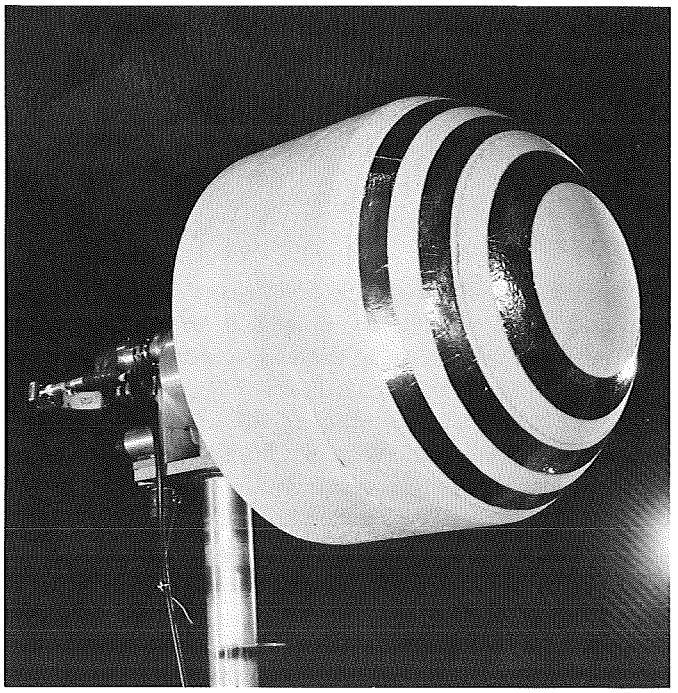


Fig. 12. Antenna in anechoic test chamber

A 2-ft-diam antenna was as large as could be easily made with the materials available. Three rings were used to provide an adequate test of the concept. The styrofoam has a low refractive index 1.015 and fills the entire space between the rings and the feed to minimize phase distortions.

Location of the foil rings on the surface is based on ray path-length calculations, taking into consideration the 60-ft distance to the illuminating horn and the refractive index of styrofoam. The ray path lengths from the illuminating horn to the waveguide phase center, at the ring edges, differ from the central ray by multiples of $\lambda/2$.

6. Experimental Results

Gain patterns for the antenna and waveguide feed are shown in Fig. 13. Peak gain is 19.5 dB with a half-power beamwidth of 3.7 deg.

For a uniformly illuminated circular aperture of this diameter, gain and beamwidth are calculated as

$$\text{Gain} = \left(\frac{\pi D}{\lambda}\right)^2 = 34.1 \text{ dB}$$

$$\text{Beamwidth} = 2 \sin^{-1}\left(0.51 \frac{\lambda}{D}\right) = 3.6 \text{ deg}$$

Efficiency is found to be

$$\eta = \frac{10^{1.95}}{10^{3.41}} = 3.5\%$$

The beamwidth is nearly the same as a uniform circular aperture, and the power lost from the main lobe is scattered into numerous sidelobes.

A phase pattern taken for the waveguide feed showed nearly constant phase up to 70 deg off axis. This verifies the assumption of a spherical wavefront with a well-defined phase center. (The edge of the outermost ring is 41 deg off the feed axis.) Gain patterns were calibrated with a 16.9-dB gain reference horn. It is estimated that the uncertainty in the measured gain is less than ± 1 dB.

7. Calculation of Losses

The overall efficiency for this antenna was found to be 3.5% as compared to 10.1% for an ideal zone plate. Allowing a ± 1 dB error on the 19.5-dB gain measurement, this efficiency could be anywhere in the range 2.8 to 4.4%.

The various factors which enter into this loss of efficiency include spillover of the feed pattern, blockage by the rings, phase errors, amplitude taper, and feed VSWR. Blockage and phase errors have already been accounted for in the calculation of the ideal case. VSWR was found to be 1.4, introducing a negligible loss of 0.12 dB. This leaves spillover and taper as the major factors.

Efficiency calculations were based on an integration of the waveguide gain pattern (as normalized by the measured gain value) to obtain total power and fraction of total power in each ring segment. A rigorous calculation would require a complete map of the gain pattern for all polar and azimuthal angles. In this case, the pattern is known for two azimuthal angles (E and H plane), and the average power for each polar angle is assumed to be the average of these two values. This assumption is strictly true only for circularly symmetric feeds but is assumed nearly correct for this case.

The integrated average power is compared to the power on axis to determine the gain of the feed. Calculated directivity for the waveguide feed is 7.6 dB, as compared to a measured gain value of 5.6 dB. The loss of 2 dB probably represents loss of power to the absorber behind the feed. The 5.6-dB gain figure was measured on two different ranges and should be accurate within ± 0.5 dB.

The efficiency of the entire aperture can be computed from a summation of the field amplitudes from each of the aperture elements (SPS 37-41, Vol. III, pp. 89–90) by the formula

$$\eta = [\Sigma (a_n P_n \eta_n)^{1/2}]^2$$

where

a_n = fractional area of nth segment

P_n = fractional power passing through nth segment

η_n = efficiency of nth segment

The fractional areas are projected areas of the rings on a plane perpendicular to the antenna axis normalized by the physical aperture as defined in Fig. 11. Fractional powers are computed by integration of the feed pattern.

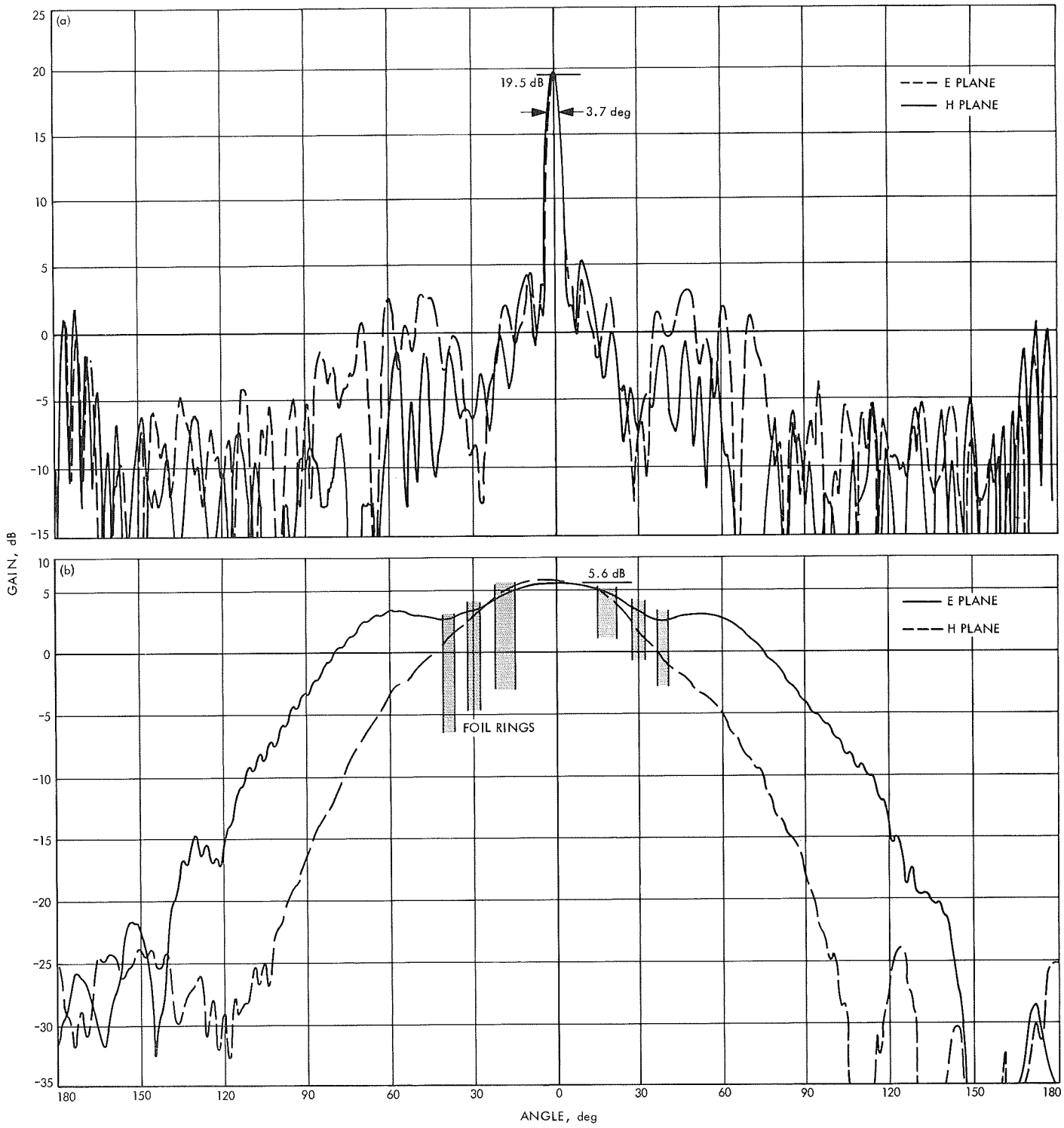


Fig. 13. Gain patterns for: (a) antenna, (b) feed

The efficiency formula can be separated into factors of spillover, phase, blockage, and taper efficiencies as follows:

$$\eta = [\Sigma (a_n P_n \eta_n)^{1/2}]^2 = [\Sigma (a_n \eta_s p_n \eta_p)^{1/2}]^2$$

$$= \eta_s \eta_p [\Sigma a_n \cdot \Sigma p_n] \cdot \frac{[\Sigma (a_n p_n)^{1/2}]^2}{\Sigma a_n \cdot \Sigma p_n}$$

$$\underbrace{\qquad}_{\eta_b} \qquad \underbrace{\qquad}_{\eta_t}$$

The spillover factor η_s has been factored out of each P_n. The segment efficiency η_n is taken as $(2/\pi)^2$ for each segment, the ideal phase efficiency for a single zone. Power included up to 41 deg is only 28% of total feed power, assuming the 5.6-dB gain figure. Spillover efficiency is thus quite low. The values obtained for a_n and P_n are found in Table 2. The overall calculated efficiency

$$\eta = \eta_s \eta_p \eta_b \eta_t = 0.0308 = 3.1\%$$

The slightly larger measured efficiency could be due to diffraction outside the third ring. The physical aperture was defined somewhat arbitrarily as having the diameter of the outside edge of the third ring. The field just outside this edge is in phase so it might be expected that the measured efficiency will be slightly higher.

Table 2. Spillover efficiency for a_n and p_n

Zone	a_n	p_n	$(a_n p_n)^{1/2}$
1	0.209	0.214	0.211
2	0.193	0.214	0.203
3	0.176	0.171	0.173
4	0.159	0.148	0.153
5	0.140	0.129	0.134
6	0.123	0.122	0.122
Σ	1.000	0.998	0.996
Open zones only			
$\Sigma a_n = 0.525$ $\Sigma p_n = 0.514$ $\Sigma (a_n p_n)^{1/2} = 0.518$			
Efficiency factors			
Spillover $\eta_s = 0.283$ Phase $\eta_p = 0.405$ Blockage $\eta_b = 0.270$ Taper $\eta_t = 0.994$			
Neglecting spillover			
$\eta_p \eta_b \eta_t = 0.109 = 10.9\%$			

8. Conclusion and Future Work

The results of this experiment indicate that a zone plate can be built on a spherical surface with no significant loss of efficiency. The overall efficiency is essentially the product of the spillover efficiency and the 10% efficiency inherent in the zone plate, and overall efficiencies of 5 to 10% should be easily achievable with better feeds. Applicability to use as a spacecraft antenna will depend on an analysis of error tolerances, a trade-off of low efficiency versus potentially larger sizes, and other problems, such as pointing losses due to the narrow beamwidth. A systems comparison with a reflector antenna should be made to determine if further work on this concept is justified. Areas for future study include: (1) evaluation of surface distortion losses which are expected to be significantly less than for a reflector antenna, (2) determination of limitations on realizable gain, and (3) investigation of surfaces other than spheres or planes.

References

1. Risser, J. R., "Dielectric and Metal-Plate Lenses," *Microwave Antenna Theory and Design*, pp. 388–412. Edited by S. Silver, Boston Technical Lithographers, M.I.T. Radiation Laboratory Series, 1963.
2. Van Buskirk, L. F., Hendrix, C. E., "The Zone Plate Lens as a Radio-Frequency Focusing Element," *IRE Trans. Ant. Prop.*, Vol. 9, No. 3, pp. 319–320, May 1961.

D. Multipactor Research: RF Voltage Breakdown in High-Impedance Coaxial Transmission Lines, R. Woo

1. Introduction

RF voltage breakdown in the 50-Ω coaxial transmission line geometry has been studied extensively in the past (Ref. 1). In this article similar studies have been carried out for higher impedance coaxial transmission lines. The purpose is to investigate the advantages of higher impedance lines as far as voltage breakdown is concerned.

2. Experimental Results

The experimental setup described in Ref. 2 was used. The experimental parameters are summarized in Table 3.

Breakdown data obtained are shown in Figs. 14 to 16. Included are data for breakdown in argon and carbon dioxide gases.

As expected, the curves are similar to those obtained for the 50-Ω coaxial geometry. The multipacting to ionization transition as well as the oscillation amplitude limit cut-off characteristics are evident in the data. For convenience of comparison the minimum power handling figures for ionization breakdown are listed in Table 4. As can be seen, the minimum breakdown voltage increases as the characteristic impedance is increased for a given value of fd (frequency times separation distance). This can be explained by the fact that as b/a is increased (b and a are the radii of outer and inner conductors, respectively), the fields become more nonuniform so that electron loss by diffusion out of the breakdown area is higher. In order to make up for this loss the breakdown voltage must be increased. Even though breakdown voltage increases with impedance, it is important to note that this is not necessarily so for breakdown power. In fact, in Table 3 it is seen that for some cases, increasing the impedance actually lowers the breakdown power. Appreciable improvement in power handling capability occurs for small values of fd, e.g., fd = 55 MHz-cm. These values correspond to cases where the electron cloud approaches the gap distance and cut-off occurs (oscillation amplitude limit). An increase in electron diffusion loss due to the nonuniformity of the fields can be

Table 3. Experimental parameters

Z_0, Ω	fd, MHz-cm	f, MHz	d, cm
74	400	106	3.77
74	200	53	3.77
74	100	26.5	3.77
74	50	13.25	3.77
91	400	72.8	5.5
91	200	36.4	5.5
91	100	18.2	5.5
91	55	10	5.5

Table 4. Minimum breakdown voltage and power for ionization breakdown

fd, MHz-cm	Z_0, ohms	Breakdown voltage, V	Breakdown power, W
400	50	72.8	106
400	74	80	86.5
400	91	98	105.3
200	50	53.5	58
200	74	66.6	60
200	91	72.6	58
100	50	49	48
100	74	57.6	45
100	91	73.9	60
55	50	57.5	66
55	74	81.5	91.5
55	91	90	92

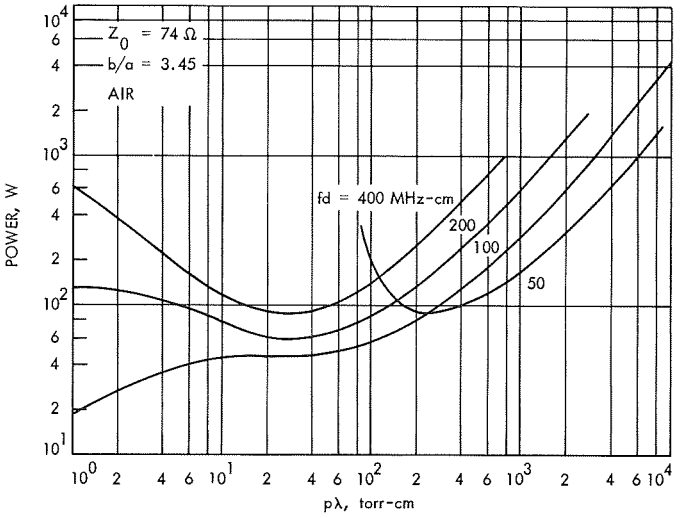


Fig. 14. Breakdown power in air versus $p\lambda$ for the coaxial transmission line geometry with characteristic impedance of 74 Ω

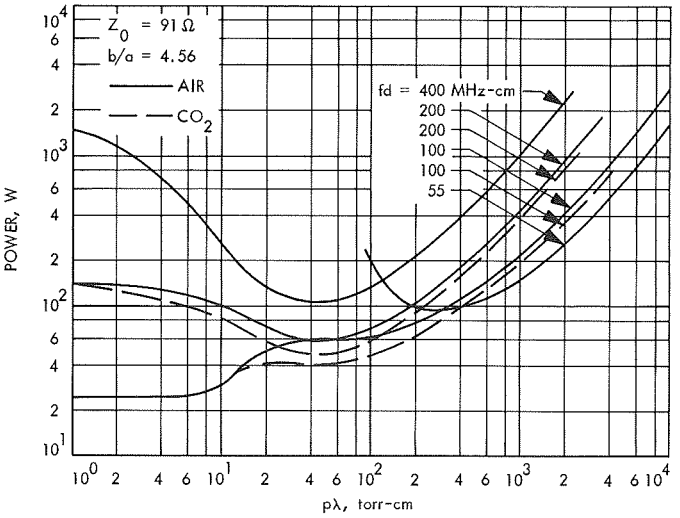


Fig. 15. Breakdown power in air and carbon dioxide versus $p\lambda$ for the coaxial transmission line geometry with characteristic impedance of 91 Ω

regarded as equivalent to a shortening of the gap distance. Since at the oscillation amplitude limit there is a rise in breakdown power as fd is decreased, improvement in power-handling capability for higher impedances would be expected. Unfortunately, for S-band frequencies, values of fd for most configurations are greater than 100 MHz-cm. For these values, there is little, if any, improvement, for higher impedances. The preceding comments apply to ionization breakdown only. For multipacting, increasing the impedance generally increases the breakdown power (Ref. 2). However, this increase is not significant. It can therefore be concluded that, in general, there is little to be gained by going to a higher characteristic impedance coaxial transmission line geometry.

It is interesting to compare the breakdown data obtained for the 50-Ω case with that obtained for higher impedances throughout the range of $p\lambda$ (pressure times wavelength). These data are shown in Fig. 17. It is seen that there is a crossover point in each case. For $p\lambda$ greater than the crossover point, less power is required for breakdown in the higher impedance case. The reverse is true for pressures less than the crossover point. This behavior can be simply explained in terms of the breakdown mechanisms involved. Above the crossover point, pressure is high enough that breakdown activity is confined to the region around the center conductor. This is confirmed visually by the extent of the breakdown glow. For this case then, breakdown is dependent on the magnitude of electric field in the area of the center conductor. Since this field increases with increasing impedance, breakdown occurs more easily in the higher impedance cases. For pressures lower than the crossover point, breakdown spreads to the outer conductor. Under these circumstances an increase in the nonuniformity of the fields results in an increase in electron loss due to diffusion to the conductor walls. In order to overcome this loss, breakdown power increases. Therefore, breakdown power is higher for higher impedances. These facts are important if the transmission line must operate in a specific pressure range. Depending on the pressure range, a higher impedance may or may not be beneficial.

3. Conclusion

Data covering RF voltage breakdown in coaxial lines with characteristic impedances higher than 50 Ω have been presented. It is seen that, generally, there is little advantage in the higher impedance lines. Many of the characteristics can be explained qualitatively. The insight gained from these studies should prove useful when studying voltage breakdown in more complicated geometries such as antennas where the fields are highly nonuniform.

References

1. Woo, R., "RF Voltage Breakdown in Coaxial Transmission Lines," *Proc. IEEE*, Vol. 57, No. 2, pp. 254–256, February 1969.

2. Woo, R., "Multipacting Discharges Between Coaxial Electrodes," *J. App. Phys.*, Vol. 39, No. 3, pp. 1528–1533, February 1968.

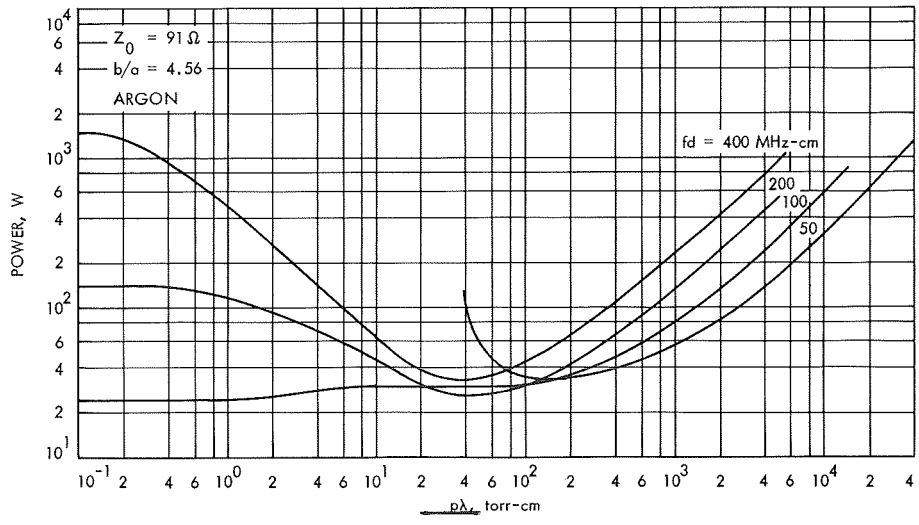


Fig. 16. Breakdown power in argon versus $p\lambda$ for the coaxial transmission line geometry with characteristic impedance of 91 Ω

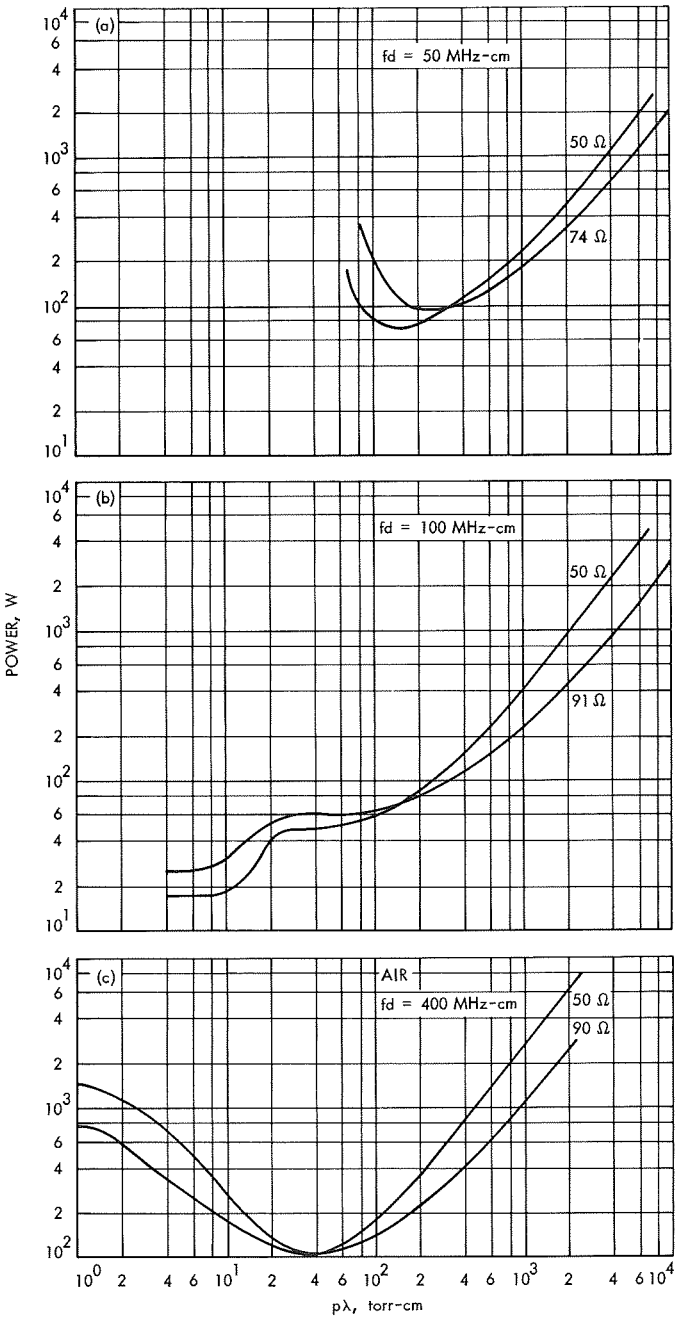


Fig. 17. Breakdown power in air versus $p\lambda$ for (a) fd = 400 MHz-cm, $Z_0 = 50\ \Omega$ and 91 Ω; (b) fd = 100 MHz-cm, $Z_0 = 50\ \Omega$ and 91 Ω; (c) fd = 55 MHz-cm, $Z_0 = 50\ \Omega$ and 74 Ω

VIII. Spacecraft Telemetry

TELECOMMUNICATIONS DIVISION

A. Efficient Spacecraft Data Systems: A Computer-Accessed Approach,

R. A. Easton and E. E. Hilbert

A powerful new approach to the Flight Telemetry Subsystem (FTS), designated as the Computer-Accessed Telemetry System (CATS) has been devised and computer simulated. Logic design of a hardware breadboard for the TOPS/Grand Tour Precursor spacecraft was begun in June. As the name implies, the traditional approach of a more-or-less fixed sampling sequence derived from a clock-controlled commutator has been scrapped in favor of being able to "set" the commutator to any desired position by loading in any desired sequence of numbers. It is no longer necessary to count "1, 2, 3, 4," etc.; and, clearly, different sampling rates can be obtained for different sensors by loading the numbers corresponding to some sensors more often than those corresponding to others. One way to do this is by using nested DO loops in a computer. The format could be changed by changing the numbers being loaded and/or their loading frequency.

Figure 1 shows the general organization of the CATS; Fig. 2 shows the conventional system for comparison. Comparison data plus information to control the format, the compression algorithm, and the allowable degree of sensor activity are stored in a memory M of approximately $M = 2(n + m) + 1$ words where n is the number of sensors and m is the number of sampling rates. ($M < 200$ for *Mariner* Mars 1971 and about 1000 for TOPS assuming roughly 500 sensors.) The information for one sensor at a time is loaded into the input/output/control (I/O/C) registers where it interacts with the control circuitry and with new data in the compressor. This results in any significant changes being read back into the memory and outputted through the storage buffers, which remove gaps due to compression and timing equalization. Two or more buffers could be used for assigning different priorities to different outputs.

If the buffers become excessively full, or if other difficulties arise, an "SOS" interrupt is sent to the central computer and sequencer (CC&S), where more sophisticated programs are stored, which will attempt to analyze and correct the problem through a format change or a change in the compression algorithm or aperture. It is important to note that the CATS does *not* require outside assistance during normal operation, or even during certain types of easy-to-diagnose abnormal operations [e.g., the activity rate of a temperature sensor exceeding that sensor's thermal time constant or an analog-to-digital converter (A/DC) failure]. The TOPS CC&S would be capable of diagnosing a problem that required correlating the outputs of a number of sensors to reach a decision. It would certainly be capable of making extensive preprogrammed format changes prior to a midcourse

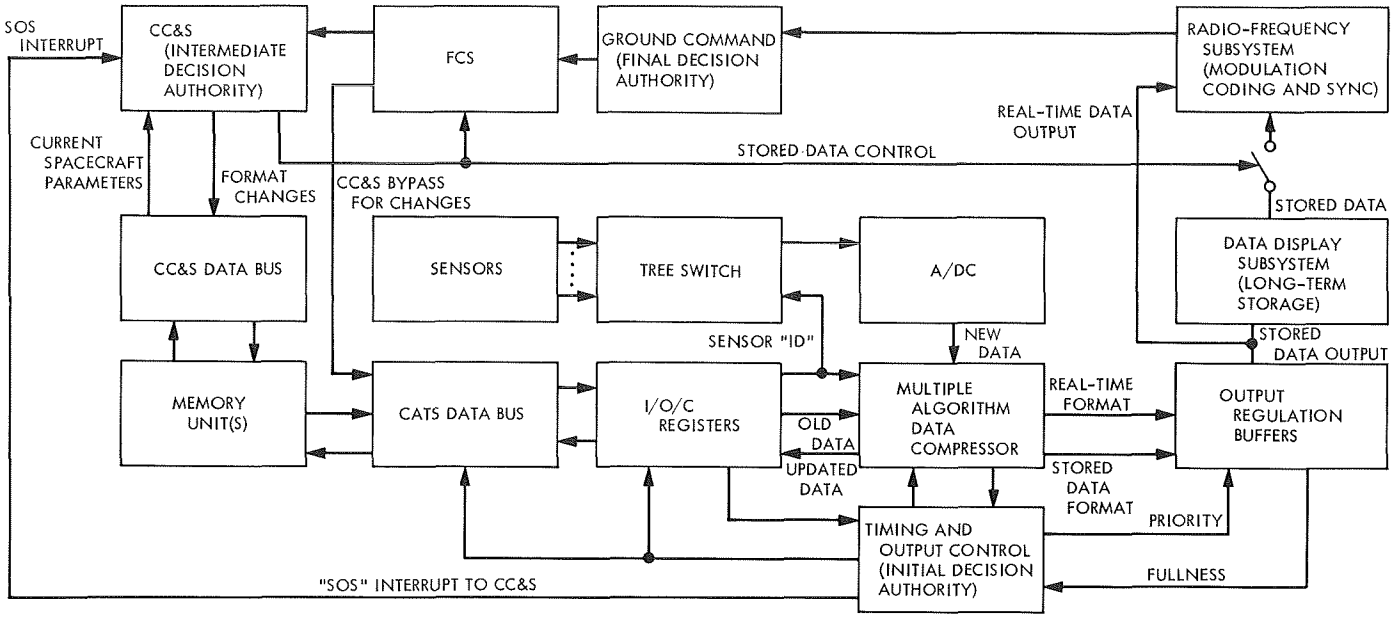


Fig. 1. CATS signal flow diagram

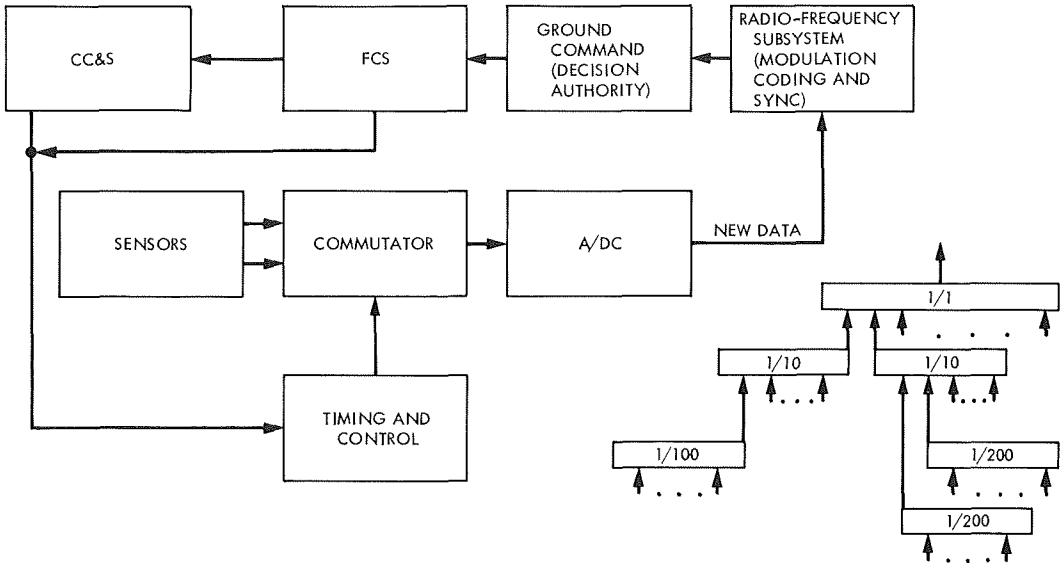


Fig. 2. Conventional FTS signal flow diagram

maneuver or other major event. The FTS can provide the CC&S with the current spacecraft data it needs in exchange for this "help," and the memory might be a shared part of the CC&S. The final decision authority rests with ground command, which can make decisions that are either too sophisticated for the CC&S or in case of CC&S failure; critical channels might be changeable only by ground command. The disadvantage of ground commands in TOPS is that the time delay could be days, or even weeks, due to the probable lack of real-time monitoring of the spacecraft during cruise, human reaction time, and the 8-h transmission delay toward the end of the mission. Therefore, a method for automating "easy-to-make" decisions is highly desirable. The key to the success of the system is an intelligent division of power between the CATS, CC&S, and ground command.

Because of the high degree of flexibility in the CATS, it possesses a gradual failure degradation character. Failure of the CC&S still allows for complete control of CATS through the Flight Command Subsystem (FCS); only the advantage of immediate control by the CC&S is lost. The loss of the FCS eliminates control by the ground user, but known required sampling format changes preprogrammed into the CC&S are still implemented. If both the FCS and CC&S fail, the CATS will continue to operate according to the sampling format directions stored in its memory. If problems occur within the CATS that cannot be corrected by the CC&S or FCS, the CATS can revert to a predecided, inflexible-type sampling format as stored in a read-only memory. This would give performance similar to that of the traditional *Mariner* system. Finally, even upon complete failure of the CATS, the entire FTS function can be taken over by the CC&S. The CC&S can provide a slightly degraded, but still flexible, FTS function per preprogrammed (and possibly updated) instructions in the CC&S memory. The CC&S is capable of providing this FTS function in a multiprocessing mode requiring a small percentage of its computation time. Thus, the CC&S can still perform its other required operations simultaneously.

The CATS itself can be made very reliable. Standby memory units are already available in the CC&S. The I/O/C registers represent mainly three shift registers of about 28 bits each that can easily be made reliable by redundancy and error-detecting coding. The timing and output control units will not be significantly more complex than their *Mariner* Mars 1969 counterparts. Critical timing logic will require redundancy. Very little hardware is needed to implement simple compressive schemes. These units could easily be made redundant, and, in addition, a failure in a specific compressive logic device could be circumvented, by simply reprogramming those sensors affected to use either an alternate compression scheme or no compression. The A/DC is a critical component and will probably be made standby redundant. The sensor-selecting tree switch is a type in which a failure in any of most of the switches causes the loss of only a few sensors. By making only a small number of the switches redundant, the entire tree can be made highly reliable. The output buffers could be made redundant by the use of primary and secondary buffers. Of course, a buffer failure would not affect the real-time mode.

Problem detection within the CATS can be easily hardware-implemented by using coding techniques developed for the STAR Computer. This type of coding will also allow identification and distinction between transient and permanent failures. Correction of transient errors is possible by providing rollback points to repeat a step involving a transient failure until it is done correctly. In the unlikely event that a transient does get through the code, it will probably only result in extra redundant samples.

Certain faults can cause the format to become mixed up, i.e., the sensors are still sampled and correctly identified, but the sampling order is wrong. However, the order can be restored by simply correcting certain key words in the memory from a redundant memory section.

Some of the other advantages of this method over the *Mariner* system are:

(1) Sampling rates can be optimized individually for each sensor.

(2) Any sensor can be added, deleted, or have its sampling rate changed without affecting the sampling rates of any other sensors. Such changes can be either automatic or ground controlled.

(3) Compression (or lack of it) can be tailored to each sensor and can be changed at will.

(4) Once a certain investment in complexity has been made, changes required to accommodate more or different sensors or sampling rates are minor. Thus, the system would be much easier to adapt to new spacecraft with "minimum change" than has been true of systems in the past.

(5) A simulation program for the CATS can accommodate any past, present, and, presumably, most future FTS formats. One use of this program would be to optimize a telemetry system on the ground. This would allow computer-aided design of conventional systems for applications where the in-flight adaptability and long-term reliability of CATS are not required.

(6) Parts of the sensor-selecting tree-switch structure can be placed in the users subsystem to allow a few wires between the systems to handle the data from a large number of sensors.

(7) Detailed information on sensor requirements is not needed until late in the project after other subsystems are fairly well defined.

(8) It is possible to sample and generate two or more simultaneous formats so that one format can be transmitted to the ground in real time while another is being used onboard the spacecraft and stored for later transmission. Either one, or both, of these channels could have data compression. Thus, both the low-rate engineering "beacon" channel and the high-rate data-dump capability of the TOPS/Grand Tour Data Handling System can be used to full advantage.

The CATS then can be regarded as a special-purpose data-handling unit providing a flexible interface between the spacecraft systems and both the onboard CC&S and ground-based command.

B. Spacecraft Command Techniques, *L. A. Couvillon*

1. Introduction

Recent work has suggested that it is possible to construct spacecraft command detection systems consisting entirely of digital logic elements; all transistors downstream of a simple analog-to-digital converter (A/DC) would be saturated or cut off. Such an approach would have a profound reliability advantage over present command systems, such as those used for the *Mariner* missions, which involve a large proportion of accurate linear circuitry, e.g., bandpass and lowpass filters. An all-digital approach is also amenable to advanced packaging technologies such as large-scale integration (LSI). This article will describe the system and logic design of one all-digital realization of a spacecraft command detector. This system is predicted to operate at pulse-code-modulated (PCM) bit rates of from 4 to 64 bits/s, at a normalized signal-to-noise ratio (SNR) of 10 dB, with bit and acquisition error rates of less than 10^{-5}.

2. Command Systems

Deep-space command systems have converged on binary PCM transmission because of its advantages both in communication efficiency and in data handling. To solve the fundamental problem of PCM bit synchronization, the *Mariner* spacecraft have used a transmitted reference scheme in which modulation power is shared between a data subcarrier and a "sync" subcarrier modulated by a pseudo-noise (PN) code (Ref. 1). Alternate methods of deriving synchronization directly from the data subcarrier have been described[1] that offer the best ultimate performance because the power allocated to the data is maximized. These improved systems have been implemented with analog circuitry.

Recent discussions[2] have led to preliminary designs for all-digital implementations of a detector that develops synchronization from the data signal. This is accomplished by computing the cross-correlation of the command modulation signal with its expected forms in sequence, and adjusting the detector phases in accordance with the correlation peaks. A shift register and serial adder are used to compute the correlations.

3. The Digital System

The fundamental algorithm of the digital detector is shown in Fig. 3. The noisy RF receiver output is correlated with a reference signal for a time T_C; if the largest correlation exceeds a threshold, indicating the presence of the command signal, the phase of the reference is adjusted to correspond to the phase of the correlation peak.

This process of correlation must be done three times: once to establish subcarrier synchronism, once for bit synchronism, and once for word synchronism (the start of actual command bits). Of course, all of these correlations could be done simultaneously, but a serial approach was chosen because the required hardware is simplified. With T_C set equal to five bit times, the time required for serial establishment of the phase references turns out to be 31 bit times; a "message header" 31-bit-times long is used to initialize and synchronize the command system prior to transmission. Figure 4 shows the flow of the initialization process.

The first signal transmitted (for 15 bit times of the 31-bit-time header) is the unmodulated subcarrier. For the next 15 bit times, the subcarrier is modulated by a PN code. The detector correlates first with an unmodulated subcarrier for subcarrier sync, then (after an appropriate wait) with the PN code to establish bit sync. The last bit time of the header is the PN code inverted. The detector recognizes this with the full force of coherent detection and begins to output data, having established all synchronization.

The reason for making the correlation interval T_C equal to five bit times is so that the rate of error in the detection of exceeding the threshold, and in the selection of the proper phase, is comparable to the probability of

[1]Carl, C., *Single-Channel Command Modulation*, (to be published).

[2]Private communication between R. M. Goldstein, E. C. Posner, and L. A. Couvillon, July 1969.

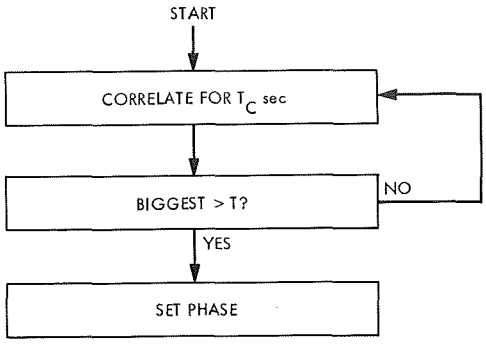


Fig. 3. Fundamental algorithm of the digital detector

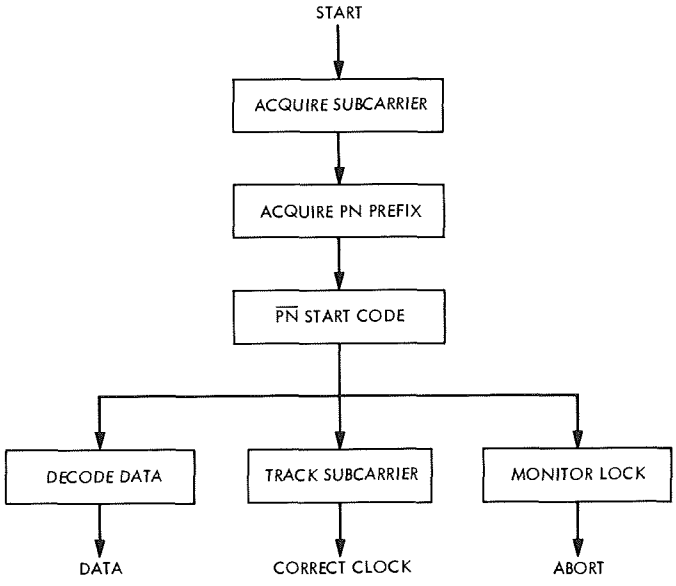


Fig. 4. Command synchronization flow

a data bit error. Allowable data error rates in a command system are typically less than 10^{-5} per bit, and making the sync decision error rates similarly small ensures virtually certain success in the acquisition and detection process. The combination of the short acquisition time (31 bits) with the high probability of acquisition (about $1–10^{-5}$) is a unique advantage of the digital system.

Figure 5 is a complete flow diagram of the digital system taking into account the three-step acquisition process, the 31 bit-time header, and the necessity to monitor both the tracking phase error in the subcarrier reference and the presence of the command signal. As shown in Fig. 5, the digital system tracks the subcarrier phase when in lock, and has a "lock detector," i.e., a test that resets the system if the absolute value of accumulated correlations fails to pass a threshold. This decision is also made with error rate 10^{-5}.

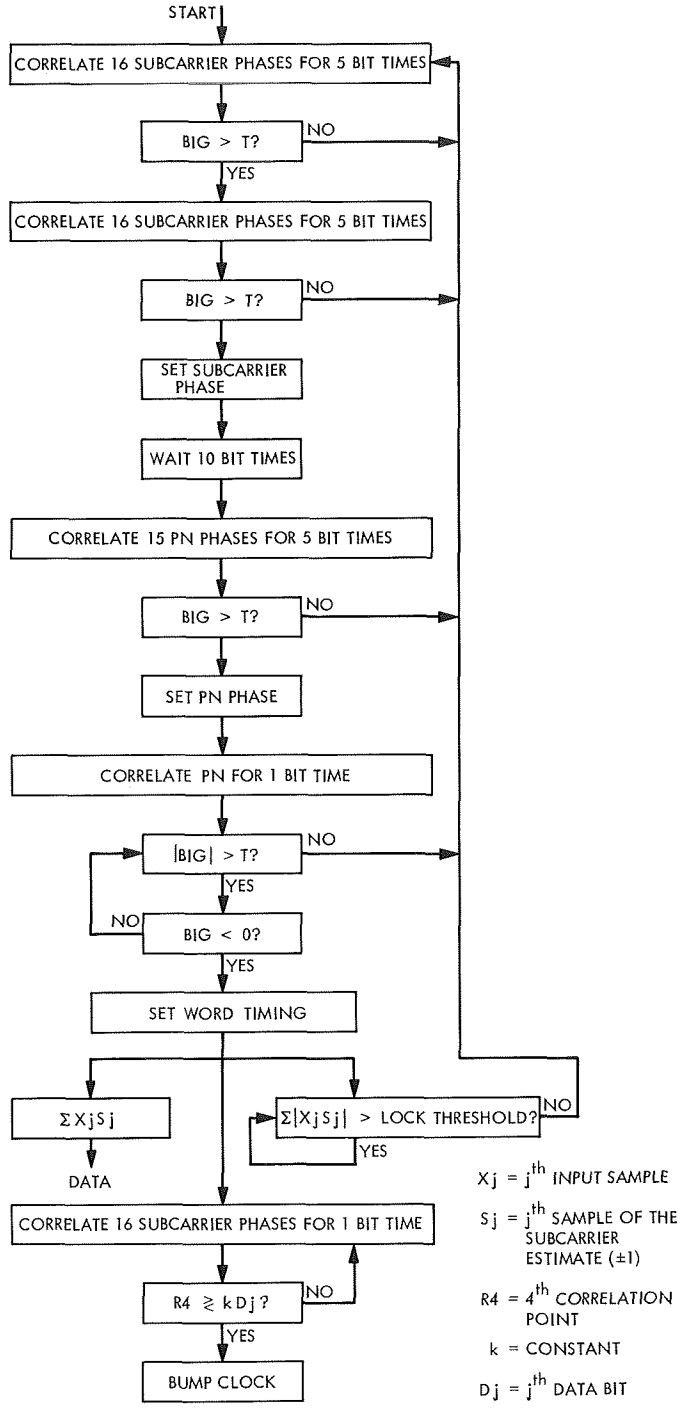


Fig. 5. Command flow diagram

4. Digital Correlation

So far, nothing has been said about how the correlation process could be implemented in spacecraft equipment. The success of the algorithm just described rests on the practicality of suitable correlation hardware. A

new implementation of a digital correlator will be used. This subsection discusses the operation of this digital correlator.

Correlation involves the computation of averaged products. However, for the case of correlation with square-wave signals having value $+1$ or -1, multiplication followed by averaging is equivalent to *addition* (or subtraction) *of new samples to an accumulated total.* The multiplier and averager in a correlator may therefore be replaced, in the case of square-wave signals, by a controllable adder/subtractor followed by an accumulator.

This concept is easily extended to the computation of a complete correlation curve by introducing multiple accumulators, one for each point to be computed on the correlation curve. Only one adder/subtractor is necessary if the whole system of accumulators can be updated during the duration of one input sample. This is shown in the diagram of Fig. 6. The input signal is the digitized value $X(t)$; the addition or subtraction is controlled by the square-wave reference X'. A system of n accumulators stores the n correlation values. All n accumulators are updated between two sample times of $X(t)$.

It is clear that this system of n accumulators can be implemented as a long shift register. Suitable registers, constructed from metal-oxide–silicon transistors in LSI circuit form, are now commercially available.

Figure 7 combines the flow diagram of Fig. 5 with the shift-register correlator of Fig. 6 into a complete all-digital command detector that is currently under construction.

Reference

1. Springett, J. C., *Command Techniques for the Remote Control of Interplanetary Spacecraft,* Technical Report 32-495, Jet Propulsion Laboratory, Pasadena, Calif., Jan. 15, 1965.

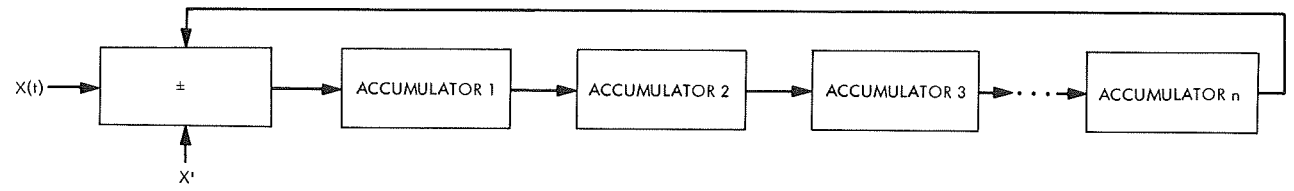


Fig. 6. Shift-register correlator

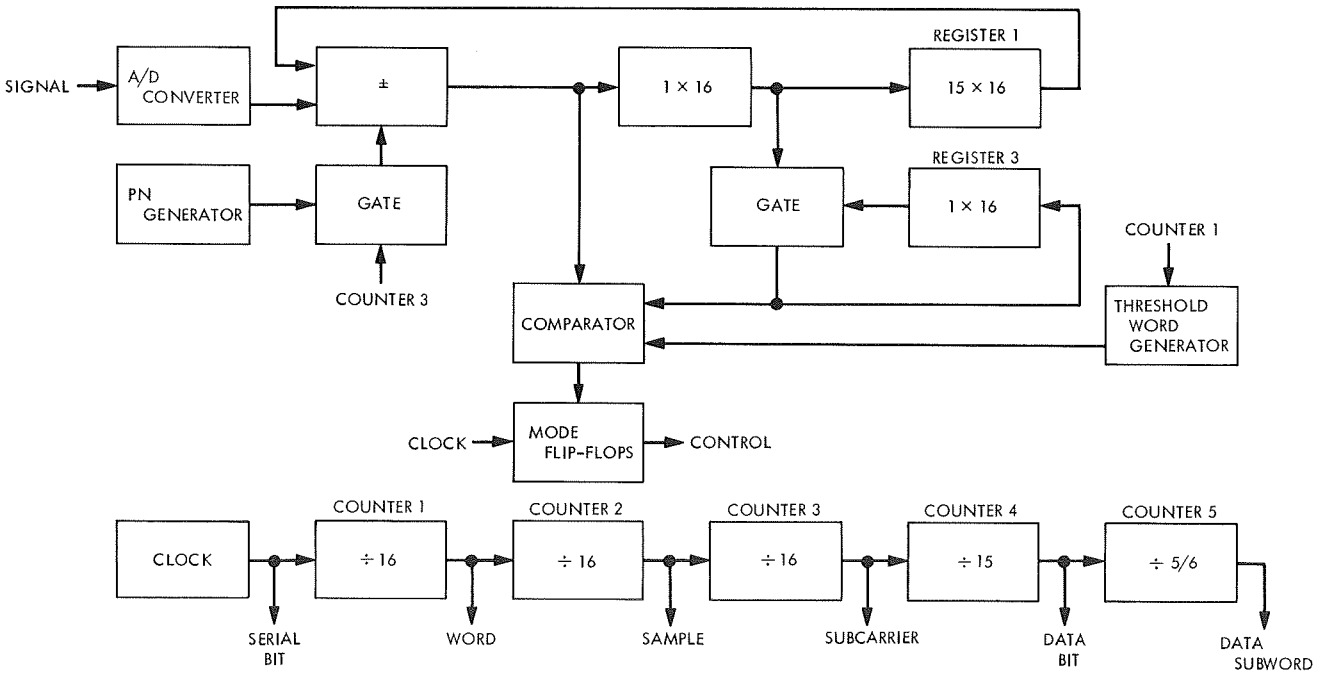


Fig. 7. Digital command system logic diagram

IX. Spacecraft Radio

TELECOMMUNICATIONS DIVISION

A. Impedance Matching With a Non-Ideal Transformer, *B. Conroy*

1. Introduction

The use of transformers for interstage coupling of solid-state amplifiers is often desirable, chiefly due to the simplicity of the circuits and their stability. Many approximate methods exist for analysis and design of transformer networks; however, for good design practice, the problem of interstage matching of solid-state power amplifiers at frequencies where core materials are not ideal requires an exact solution.

A mathematical statement of the problem is as follows: Given a circuit as shown in Fig. 1 with a generator with source resistance R_g, a load with resistance R_L, a transformer with coupling factor k, and a single capacitor for tuning the primary, find

(1) The minimum Q (corresponding to the maximum bandwidth) of the coupling network.

(2) The values of the primary reactance X_p, secondary reactance X_s, and capacitive reactance X_c to realize the minimum Q.

(3) The values of X_p, X_s, and X_c, which will give an arbitrary Q.

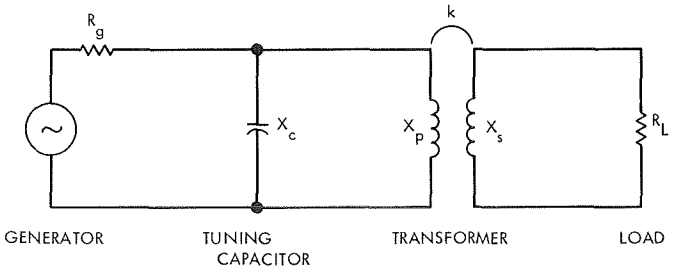


Fig. 1. Transformer-matching network

2. Basic Analysis

This circuit involves two impedance transformations. One is due to transformer action, and the other is a resonant transformation due to the presence of leakage inductance in the transformer and X_c. The transformation from X_c results in a series equivalent resistance R_s and a series equivalent capacitive reactance X_E given by

$$R_s = \frac{R_g}{1 + \left(\frac{R_g}{X_c}\right)^2} \tag{1}$$

and

$$X_E = \frac{X_c}{1 + \left(\frac{X_c}{R_g}\right)^2} \tag{2}$$

The impedance looking into the primary of the transformer will be a series equivalent resistance R'_s and a series equivalent inductive reactance X'_E given by (Ref. 1)

$$R'_s = \frac{k^2 R_L X_p X_s}{R_L^2 + X_s^2} \tag{3}$$

and

$$X'_E = X_p - \frac{k^2 X_p X_s^2}{R_L^2 + X_s^2} \tag{4}$$

Since X_E is capacitive and X'_E is inductive, the load will be matched to the generator if and only if

$$X_E = X'_E \tag{5}$$

and

$$R_s = R'_s \tag{6}$$

This results in

$$R_s = \frac{k^2 R_L X_p X_s}{R_L^2 + X_s^2} \tag{7}$$

and

$$X_E = X_p - \frac{k^2 X_p X_s^2}{R_L^2 + X_s^2} \tag{8}$$

Solving Eq. (7) for X_p,

$$X_p = \frac{R_s (R_L^2 + X_s^2)}{k^2 R_L X_s} \tag{9}$$

and from Eq. (8),

$$X_p = \frac{X_E (R_L^2 + X_s^2)}{R_L^2 + X_s^2 - k^2 X_s^2} \tag{10}$$

Equating these results,

$$\frac{R_s (R_L^2 + X_s^2)}{k^2 R_L X_s} = \frac{X_E (R_L^2 + X_s^2)}{R_L^2 + X_s^2 (1 - k^2)} \tag{11}$$

and simplifying,

$$R_s R_2^L + R_s X_s^2 (1 - k^2) = k^2 R_L X_s X_E \tag{12}$$

or

$$X_s^2 [R_s (1 - k^2)] - X_s [k^2 X_E R_L] + R_s R_L^2 = 0 \tag{13}$$

This may be solved with the quadratic formula as

$$X_s = \frac{k^2 X_E R_L \pm [k^4 X_E^2 R_L^2 - 4 R_s^2 R_L^2 (1 - k^2)]^{1/2}}{2 R_s (1 - k^2)} \tag{14}$$

or

$$X_s = \frac{R_L}{2 R_s (1 - k^2)} \left\{ k^2 X_E \pm [k^4 X_E^2 - 4 R_s^2 (1 - k^2)]^{1/2} \right\} \tag{15}$$

For there to be real solutions for X_s, the inequality

$$k^4 X_E^2 \geq 4 R_s^2 (1 - k^2) \tag{16}$$

must be satisfied, or, in different form,

$$\left(\frac{X_E}{R_s}\right)^2 \geq \frac{4 (1 - k^2)}{k^4} \tag{17}$$

In the matched condition,

$$Q = \frac{X_E}{2 R_s} \tag{18}$$

so the inequality may be written as

$$Q^2 \geq \frac{1 - k^2}{k^4} \tag{19}$$

3. Minimum Q Solution

Equation (19) implies that the minimum Q is

$$Q_{\min} = \frac{(1 - k^2)^{1/2}}{k^2} \tag{20}$$

and this corresponds to the case

$$\left(\frac{X_E}{R_s}\right)^2 = \frac{4 (1 - k^2)}{k^4} \tag{21}$$

From Eqs. (1) and (2),

$$\left(\frac{R_s}{X_E}\right)^2 = \left(\frac{R_g}{X_c}\right)^2 \left\{\frac{1+\left(\frac{X_c}{R_g}\right)^2}{1+\left(\frac{R_g}{X_c}\right)^2}\right\}^2 \tag{22}$$

which reduces to

$$\left(\frac{R_s}{X_E}\right)^2 = \left(\frac{R_g}{X_c}\right)^2 \left(\frac{X_c}{R_g}\right)^4 \left(\frac{R_g^2+X_c^2}{X_c^2+R_g^2}\right)^2 \tag{23}$$

or

$$\left(\frac{R_s}{X_E}\right)^2 = \left(\frac{X_c}{R_g}\right)^2 \tag{24}$$

These results can be substituted into Eq. (1) to give

$$R_s = \frac{R_g}{1+\frac{4(1-k^2)}{k^4}} \tag{25}$$

Rewriting Eq. (21),

$$X_E = \frac{2R_s(1-k^2)^{1/2}}{k^2} \tag{26}$$

and combining Eqs. (21) and (24),

$$X_c = \frac{k^2 R_g}{2(1-k^2)^{1/2}} \tag{27}$$

Substituting Eqs. (24) and (26) into Eq. (15),

$$X_s = \frac{R_L}{2R_s(1-k^2)}\left\{k^2 R_s\left[\frac{4(1-k^2)}{k^4}\right]^{1/2}\right\} \tag{28}$$

which reduces to

$$X_s = \frac{R_L}{(1-k^2)^{1/2}} \tag{29}$$

Substitution of the results of Eqs. (25) and (29) into Eq. (9) produces

$$X_p = \frac{R_g}{k^2}\left[\frac{k^4}{k^4+4(1-k^2)}\right]\left[(1-k^2)^{1/2}+\frac{1}{(1-k^2)^{1/2}}\right] \tag{30}$$

or

$$X_p = \frac{k^2 R_g}{(1-k^2)^{1/2}(2-k^2)} \tag{31}$$

4. Arbitrary Q Solution

In the general case, Eqs. (18) and (24) are still true, so

$$Q = \frac{X_E}{2R_s} \tag{32}$$

and

$$Q = \frac{R_g}{2X_c} \tag{33}$$

Thus,

$$X_c = \frac{R_g}{2Q} \tag{34}$$

Substitution of Eqs. (20) and (32) into Eq. (15) gives

$$X_s = \frac{R_L}{(1-k^2)^{1/2}}\left(\frac{Q}{Q_{\min}}\right)\left\{1 \pm \left[1-\left(\frac{Q_{\min}}{Q}\right)^2\right]^{1/2}\right\} \tag{35}$$

and substitution of Eqs. (1) and (33) into Eq. (9) gives

$$X_p = \frac{R_g}{k^2(1+4Q^2)}\left[\frac{R_L}{X_s}+\frac{X_s}{R_L}\right] \tag{36}$$

5. Concluding Remarks

In a practical design problem, the first step is to find the minimum Q (Eq. 20). Then, if this Q is acceptable for the desired network, X_c, X_s, and X_p are found from Eqs. (27), (29), and (31), respectively.

If a Q greater than $Q_{\min}$ is desired, X_c is found from Eq. (34). Two possible values of X_s are then found from Eq. (35) and one of them is chosen. Finally, the value of X_p corresponding to the chosen value of X_s is found from Eq. (36).

Reference

1. Terman, F. E., *Electronic and Radio Engineering*, 4th Edition, pp. 57–62. McGraw-Hill Book Co., Inc., New York, 1955.

B. Distributed Amplifier Evaluation, *B. Conroy*

1. Introduction

Since the power-handling capability of high-frequency transistors is limited, one key to their successful use in transmitter applications is to combine several of these transistors to increase the available output power. Of the several methods of power combinations available, one that seemed promising, but inadequately documented, was a distributed configuration in which amplifier stages were placed between input and output transmission lines.

A literature search revealed only one reference to a high-power, distributed amplifier. This work was done by RCA, Princeton, under contract to the Air Force Avionics Laboratory; the goal was to produce 300 W at 400 MHz. Work on the distributed configuration was terminated, however, due to a change in contract scope calling for octave bandwidth. Data from this work indicated that, while power output and efficiency were not impaired, overall gain was less than that of an individual stage. However, it was felt that, by changing input line impedance, the problem might be overcome. On the basis of this information, it was decided to investigate this method of power combination.

Two versions (each with an 860-MHz nominal frequency) were designed and tested. The first version (stripline), which used United Aircraft S-1050 stripline transistors, is shown in Fig. 2; the second version (lumped-constant), which used RCA 2N5470 coaxial transistors, is shown in Fig. 3.

2. Design Considerations

The fact that efficiency and power gain are the main objectives of a high-power, solid-state distributed amplifier necessitates a conceptual variation from the classical vacuum tube unit where voltage gain and wide bandwidth are the major objectives. Figure 4 is a block diagram of the high-power, solid-state distributed amplifier with four amplifier stages. The difference between this and the classical version is the absence of terminations on the unused ends of the input and output lines. The resulting standing waves on the lines must be matched by making the lines resonant and equating input and output impedances of the amplifier stages with the number of stages times the generator and load impedances.

The individual amplifier stages were designed using standard RF design techniques. Figure 2b shows the stripline amplifier with its cover removed. There are two stubs on each base. Each collector has two tuning stubs plus an extra quarter wave stub, which is terminated in a capacitor, to provide low impedance for the second harmonic frequency and increase efficiency. In the first stage, each stub has been shortened and loaded with a variable capacitor to provide some adjustment capability.

Figure 5 is a schematic of an amplifier section of the lumped-constant amplifier. This is essentially the test circuit used by RCA for these transistors.

3. Experimental Procedure and Data

In testing the stripline distributed amplifier, several difficulties were encountered: (1) to do any tuning, the cover had to be removed and the lines trimmed, (2) instability of the S-1050 transistor, and (3) difficulty in measuring the performance of each amplifier stage.

Since experience had shown the 2N5470 to be a more stable transistor, the stripline amplifier was modified to accept this device. However, this modification proved to be inadequate; because of parameter differences and the inductance added by the wires connecting the coaxial transistor to the stripline, the amplifier would not tune, even after the tuning stubs had been modified and variable capacitors had been added. It was determined that a new design would be required for satisfactory performance with the coaxial transistors. Since it was desired to evaluate the distributed configuration rather than the specific circuit medium, and due to the other difficulties referred to above, the stripline version of the distributed amplifier was abandoned.

The lumped-constant model was then built. With this arrangement, it was possible to test each individual stage and the complete distributed amplifier with different impedance input and output lines.

Table 1 presents the data using 12.5- and 50-Ω input and output lines. The tuning was quite critical, and equal sharing of the load was almost impossible to obtain.

4. Conclusions

In spite of its fewer parts and freedom from power combiner or isolator loss, the distributed amplifier can be designated as impractical for high-power, solid-state amplifiers. No simple, reliable solutions could be found for its inherent difficulties:

(1) Direct feedback from each amplifier to all others.

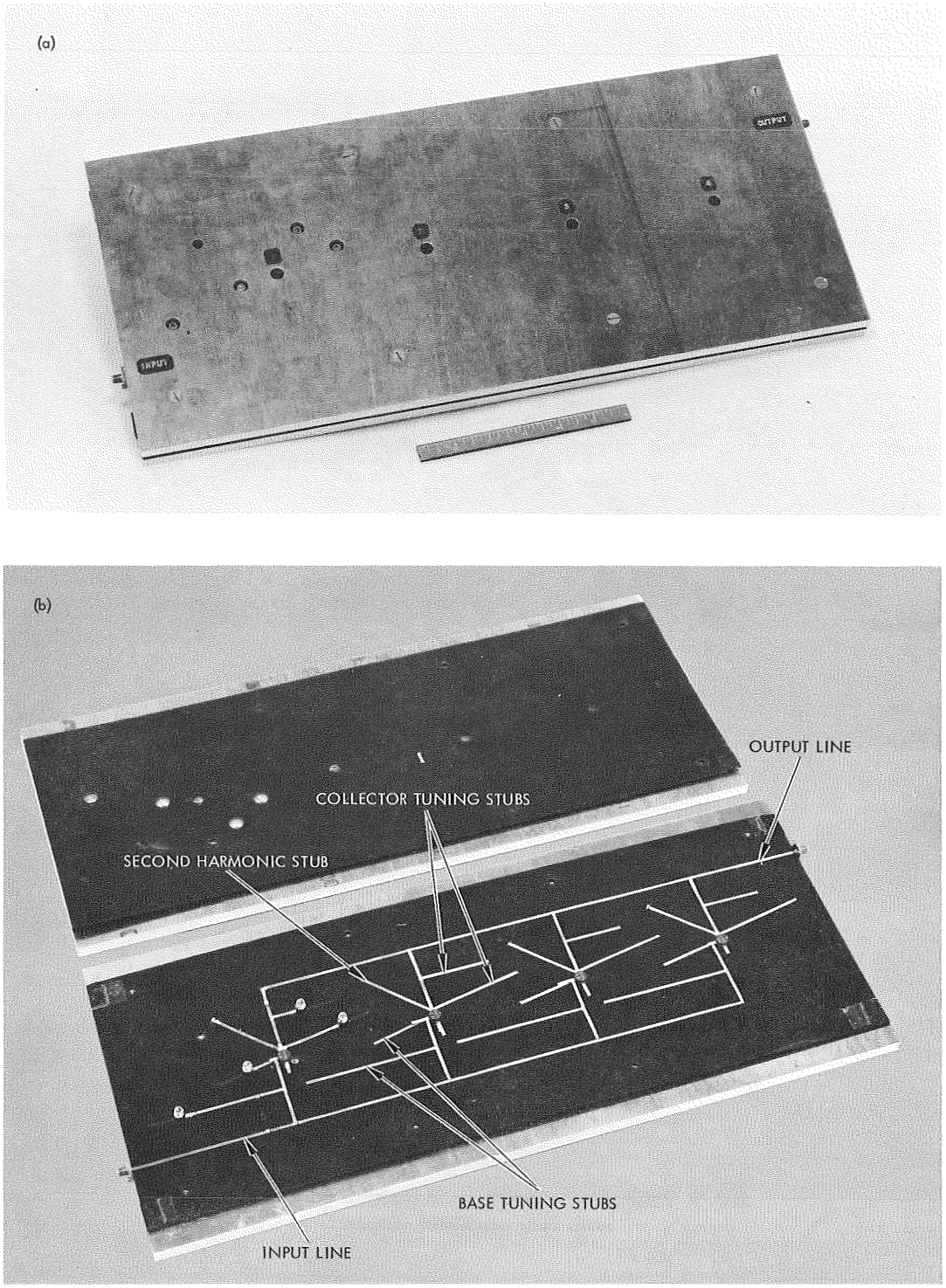


Fig. 2. Stripline distributed amplifier: (a) assembled, (b) cover removed

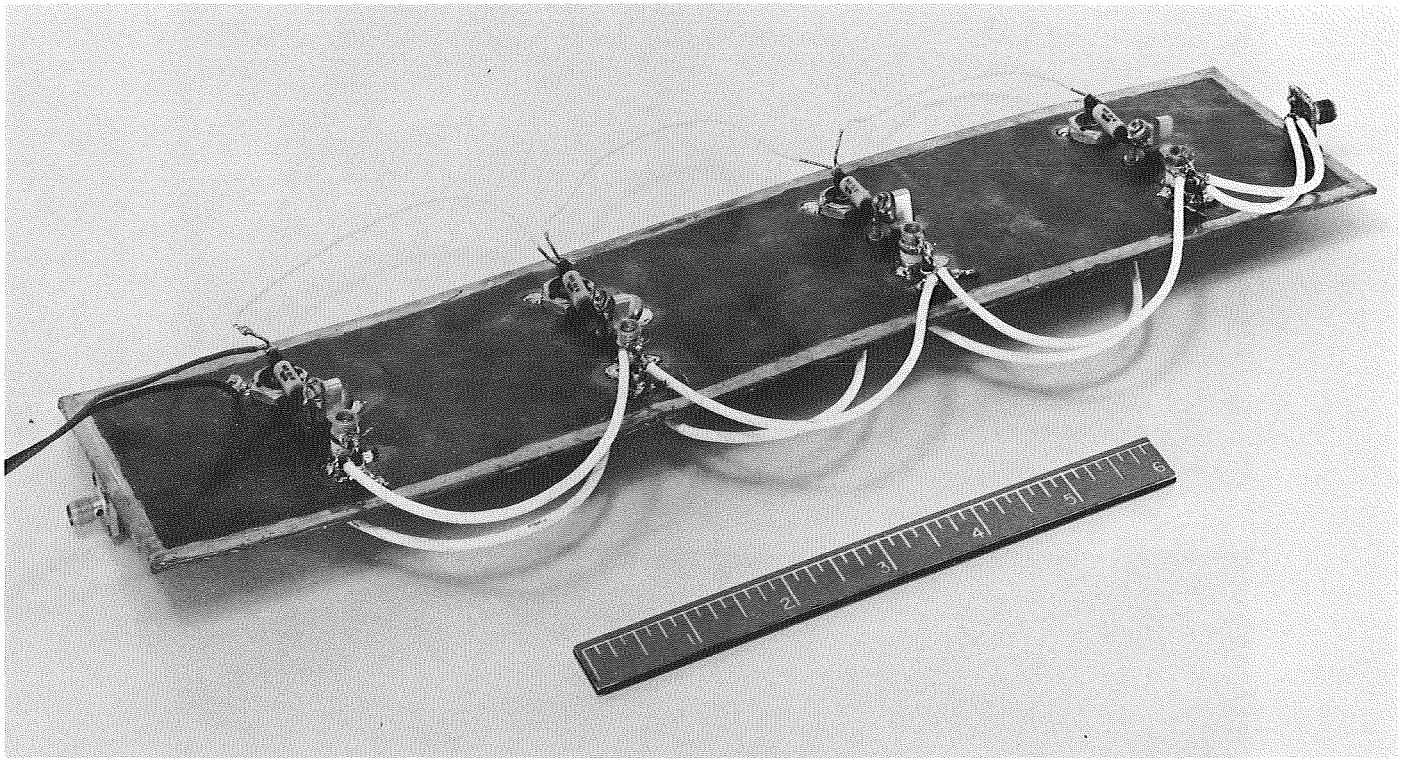


Fig. 3. Lumped-constant distributed amplifier

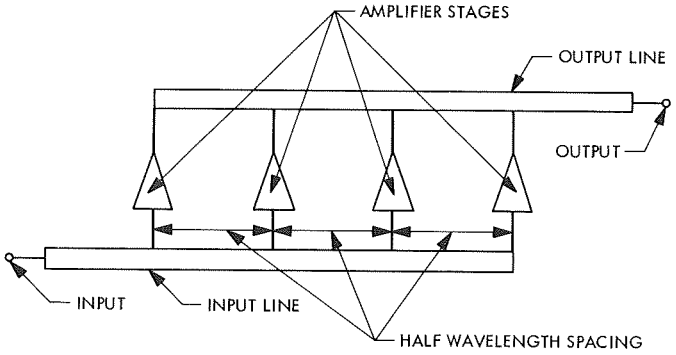


Fig. 4. High-power, solid-state distributed amplifier with four amplifier stages

(2) No isolation to compensate for unbalances in the power distribution, stage gain, and power combination.

Table 1. Performance data on distributed amplifier using 2N5470

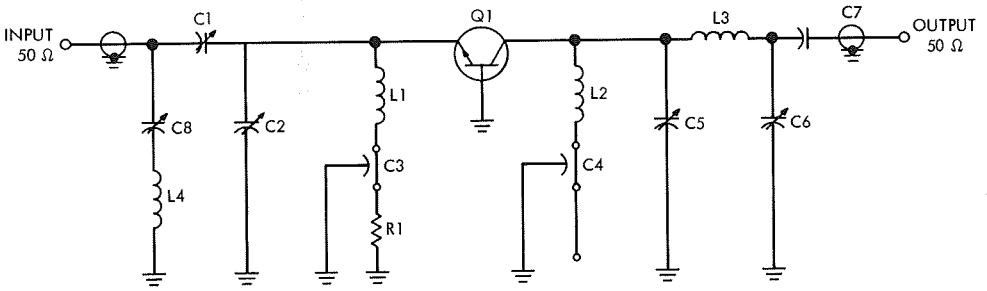
Test item	Gain, dB	Power output, W	Efficiency, %
Individual amplifier stages	6 to 6.3	1 to 1.2	60 to 63
Complete amplifier with 12.5-Ω input and output lines	4.8	3.2	60
Complete amplifier with 50-Ω input and output lines	4.6	3.1	59

Correction of the first difficulty requires use of extremely stable transistors (i.e., those with low internal feedback at all frequencies) and the second requires close control of input and output circuits and individual stages.

In view of the experimental evidence, it has become apparent that the hybrid tree, Wilkinson Summer (Ref. 1), or ferrite isolator/power combiner is more practical for spacecraft transmitter applications than the envisioned distributed amplifier configuration.

Reference

1. Wilkinson, E. J., "An N-Way Hybrid Power Divider," *IRE Transactions on Microwave Theory and Techniques*, pp. 116–118, Jan. 1960.



L3 0.01 × 0.157-in. SILVER STRIP 0.90-in. LONG
L4 0.025-in.-diam SILVER WIRE 0.50-in. LONG
Q1 2N5470 RCA COAXIAL TRANSISTOR
R1 20 Ω 0.5 W CARBON

C1, C2, C5, C6, C8 0.3 to 10 pF
C3, C4 1000 pF FEEDTHROUGH
C7 1000 pF LEADLESS CERAMIC
L1, L2 0.1 μH RF CHOKE

Fig. 5. Lumped-constant amplifier section

X. Spacecraft Communication Systems

TELECOMMUNICATIONS DIVISION

A. Fast Two-Dimensional Fourier Transform,

J. P. Hong

1. Introduction

The direct discrete one-dimensional Fourier transform of N points requires N^2 operations (complex multiplication followed by addition). If the N points are samples of a signal spaced T/N sec, the transform gives $N/2$ point pairs spaced $1/T$ Hz. Based on the interaction algorithm of t^n factorial experiments, Good (Ref. 1) reported a Fourier transform algorithm which requires NQ_N operations. Q_N is a function of the mutually prime factors of N. For the important case when $N = 2^n$ this reduces to $N \log_2 N$. Cooley and Tukey (Refs. 2 and 3) developed a similar algorithm for high-speed computers. Cochran, et al., (Ref. 4) pointed out that certain multiplication need not be performed and reduced the number of complex multiplications to $N/2 \log_2 N$. Andrews (Ref. 5) further eliminated redundant computations and showed an algorithm which requires $N \log N$ complex additions and $N/2\ [\log_2 (N)-2] + 1$ complex multiplications.

The algorithms for two-dimensional Fourier transform proposed by Good and Cooley–Tukey are essentially the direct application of the one-dimensional fast Fourier transform algorithm to the $N \times N$ points. The number of operations required for the transform as reported by these authors is

$$N^2 \log_2 N^2 = NN \log_2(N \cdot N) = 2N[N \log_2 N]$$

It is pointed out by Cochran (Ref. 4) and Pratt and Andrews (Ref. 6) that the kernel of the Fourier transform is separable (Eq. 3). They propose the repeated use of the one-dimensional transform (Eq. 1) to obtain a two-dimensional transform. Each row is first transformed (Eq. 4); then the transform is applied to each column (Eq. 5). This procedure does not necessarily produce an optimum algorithm. The number of complex multiplications and complex additions required by each method is outlined in Table 1.

$$F(u,y) = \mathcal{F}_x[f(x,y)] = \sum_{x=0}^{N-1} W(ux/N) f(x,y) \tag{1}$$

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) W((ux+vy)/N) \tag{2}$$

$$= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) W(ux) \exp(-2\pi jvy) \tag{3}$$

$$= \sum_{y=0}^{N-1} \exp(-2\pi jvy/N) \left\{ \sum_{x=0}^{N-1} W(ux/N) f(x,y) \right\} \tag{4}$$

$$= \sum_{y=0}^{N-1} W(vy/N) F(u,y) \tag{5}$$

$$= \mathcal{F}_y \{ \mathcal{F}_x[f(x,y)] \} \tag{6}$$

Table 1. Complex multiplications and additions

Method	One-dimensional		Two-dimensional	
	Additions[a]	Multiplications[a]	Additions[a]	Multiplications[a]
Direct	N^2	N^2	$2N^3$	$2N^3$
Cooley–Tukey	$N \log_2 N$	$N \log_2 N$	$2N^2 \log_2 N$	$2N^2 \log_2 N$
Cochran	$N \log_2 N$	$\frac{1}{2} N \log_2 N$	$2N^2 \log_2 N$	$N^2 \log_2 N$
Andrews	$N \log_2 N$	$\frac{1}{2} N [\log_2 (N) - 2] + 1$	$2N^2 \log_2 N$	$N^2 [\log_2 (N) - 2] + 2N$
Hong			$2N^2 \log_2 N$	$\frac{3}{4} N^2 [\log_2 (N) - 1]$

[a]Complex additions and complex multiplications.

2. Decimation in the Two-Dimensional Sample Domain

Suppose $f(x,y)$ is a matrix of sample points. Let us consider a two-dimensional Fourier transform of it:

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \exp\left(-2\pi j(ux+vy)/N\right) \tag{7}$$

We introduce the notation

$$\begin{aligned} f_0(p,q) &= f(2p,2q) \qquad f_1(p,q) = f(2p+1,2q) \qquad 0 \le p < N/2 \\ f_2(p,q) &= f(2p,2q+1) \qquad f_3(p,q) = f(2p+1,2q+1) \qquad 0 \le q < N/2 \end{aligned} \tag{8}$$

Using Eq. (8), we can rewrite Eq. (7)

$$\begin{aligned} F(u,v) = &\sum_{x=\text{even}} \sum_{y=\text{even}} f(x,y) W\left((ux+vy)/N\right) \\ &+ \sum_{x=\text{odd}} \sum_{y=\text{even}} f(x,y) W\left((ux+vy)/N\right) \\ &+ \sum_{x=\text{even}} \sum_{y=\text{odd}} f(x,y) W\left((ux+vy)/N\right) \\ &+ \sum_{x=\text{odd}} \sum_{y=\text{odd}} f(x,y) W\left((ux+vy)/N\right) \end{aligned} \tag{9}$$

$$\begin{aligned} F(u,v) = &\sum_{p=0}^{p<N/2} \sum_{q=0}^{q<N/2} f_0(p,q) W[(up+vq)/(N/2)] \\ &+ \sum_{p=0}^{p<N/2} \sum_{q=0}^{q<N/2} f_1(p,q) W[(up+vq)/(N/2)] W[u/N] \\ &+ \sum_{p=0}^{p<N/2} \sum_{q=0}^{q<N/2} f_2(p,q) W[(up+vq)/(N/2)] W[v/N] \\ &+ \sum_{p=0}^{p<N/2} \sum_{q=0}^{q<N/2} f_3(p,q) W[(up+vq)/(N/2)] W[(u+v)/N] \end{aligned}$$

We use the notation $F_1(u,v) = \mathcal{F}^{(2)}[f_1(x,y)]$, which is periodic in u and v of period $N/2$. Therefore, we can rewrite Eq. (10), and obtain Eq. (11a).

We make use of an important equality Eq. (12) which allows us to complete the transform.

That is, $F(u,v)$, $F(u+N/2,v)$, $F(u,v+N/2)$, and $F(u+N/2,v+N/2)$ are all very closely related and computation of one essentially yields all of them.

$$F(u,v) = F_0(u,v) + F_1(u,v)W(u/N) + F_2(u,v)W(v/N) + F_3(u,v)W((u+v)/N) \qquad 0 \le u < N/2, \quad 0 \le v < N/2 \tag{11a}$$

$$F(u+N/2,v) = F_0(u,v) - F_1(u,v)W(u/N) + F_2(u,v)W(v/N) - F_3(u,v)W((u+v)/N) \tag{11b}$$

$$F(u,v+N/2) = F_0(u,v) + F_1(u,v)W(u/N) - F_2(u,v)W(v/N) - F_3(u,v)W[(u+v)/N] \tag{11c}$$

$$F(u+N/2,v+N/2) = F_0(u,v) - F_1(u,v)\,W(u/N) - F_2(u,v)W(v/N) + F_3(u,v)W[(u+v)/N] \tag{11d}$$

$$W[(p+N/2)/N] = W(p/N)W[(N/2)/N] = -W(p/N) \tag{12}$$

If we have $F_0(u,v)$, $F_1(u,v)$, $F_2(u,v)$, $F_3(u,v)$, which are each $(N/2) \times (N/2)$, then we can compute $F(u,v)$ which is $N \times N$ in $0.75N^2$ multiplications and $2N^2$ additions, since three multiplications (Eq. 13) for each of $(N/2) \times (N/2)$ points and eight additions (Eqs. 14 and 15) give four of the $N \times N$ transform points (Eq. 16).

$$\left.\begin{aligned} a &= F_0(u,v) \\ b &= F_1(u,v)W(u/N) \\ c &= F_2(u,v)W(v/N) \\ d &= F_3(u,v)W(u+v)/N) \end{aligned}\right\} \tag{13}$$

$$\left.\begin{aligned} e &= a + c \\ f &= b + d \\ g &= a - c \\ h &= b - d \end{aligned}\right\} \tag{14}$$

$$\left.\begin{aligned} F(u,v) &= e + f \\ F(u+N/2,v) &= e - f \\ F(u,v+N/2) &= g + h \\ F(u+N/2,v+N/2) &= g - h \end{aligned}\right\} \qquad \begin{aligned} 0 \le u < N/2 \\ 0 \le v < N/2 \end{aligned} \tag{15}$$

Let us consider a two by two matrix. We note that

$$\begin{aligned} F_0 &= f(0,0) & F_1 &= f(1,0) \\ F_2 &= f(0,1) & F_3 &= f(1,1) \end{aligned} \tag{16}$$

Therefore, using Eqs. (13), (14), and (15), we obtain all of the transform points $F(0,0)$, $F(1,0)$, $F(0,1)$, $F(1,1)$:

$$\left.\begin{aligned} a &= f(0,0) \\ b &= f(1,0) \\ c &= f(0,1) \\ d &= f(1,1) \\ e &= f(0,0) + f(0,1) \\ f &= f(1,0) + f(1,1) \\ g &= f(0,0) - f(0,1) \\ h &= f(1,0) - f(1,1) \\ F(0,0) &= f(0,0) + f(0,1) + f(1,0) + f(1,1) \\ F(1,0) &= f(0,0) + f(0,1) - f(1,0) - f(1,1) \\ F(0,1) &= f(0,0) - f(0,1) + f(1,0) - f(1,1) \\ F(1,1) &= f(0,0) - f(0,1) - f(1,0) + f(1,1) \end{aligned}\right\} \tag{17}$$

In a two-dimensional Fourier transform a two-by-two matrix is a primitive set. That is, it is the largest matrix whose transform can be found without multiplication.

The cost of deriving the transform of an N-by-N matrix from four appropriate sets of transforms, each $N/2$ by $N/2$, is $8(N/2)(N/2) = 2N^2$ additions and $3(N/2)(N/2) = 3N^2/4$ multiplications. The total number of additions which are required to find an N by N transform from

the primitive sets is $2N^2 \log_2 N$ complex additions and $3N^2/4(-1 + \log N)$ complex multiplications (Fig. 1):

$$\text{Number of additions} = \frac{8}{4}\left\{N^2 + 4\left[\left(\frac{N}{2}\right)^2 + 4\left[\left(\frac{N}{4}\right)^2 + \cdots 4[(2)^2]\cdots\right]\right]\right\}$$

$$= 2\{(2^n)^2 + 2^2\left[(2^{n-1})^2 + 2^2\left[(2^{n-2})^2 + \cdots 2^2[(2^1)^2]\cdots\right]\right]\}$$

$$= 2N^2 \log N \tag{18}$$

$$\text{Number of multiplications} = \frac{3}{4}\{(2^n)^2 + 2^2\left[(2^{n-1})^2 + 2^2\left[(2^{n-2})^2 + \cdots 2^2\left[(2^2)^2 \cdots\right]\right]\right]\} \tag{19}$$

$$= \frac{3}{4}N^2(-1 + \log_2 N) \qquad N \geq 2$$

3. Conclusion

Whereas Andrews' method is faster than all earlier algorithms, the ratio of effort needed by his algorithm to that of the others approaches unity as $N = 2^n$ becomes large:

$$\lim_{n\to\infty}\frac{N^2[\log_2(N)-2] + 2N}{N^2\log_2 N} = \lim_{n\to\infty}\frac{N^2\log_2 N}{N^2\log_2 N} - \frac{2N^2}{N^2\log_2 N} + \frac{2N}{N^2\log_2 N} = 1 \tag{20}$$

It is immediately obvious that the author's algorithm gives faster transforms for nontrivial N and that the ratios (Eqs. 21 and 22) approach 0.75 as N becomes large.

Comparison with Cochran (Figs. 2 and 3) shows

$$\lim_{N\to\infty}\frac{\frac{3}{4}N^2[\log_2(N)-1]}{N^2\log_2 N} = \lim_{N\to\infty}\frac{3}{4}\frac{N^2\log_2 N}{N^2\log_2 N} - \frac{\frac{3}{4}N^2}{N^2\log N} = \frac{3}{4} \tag{21}$$

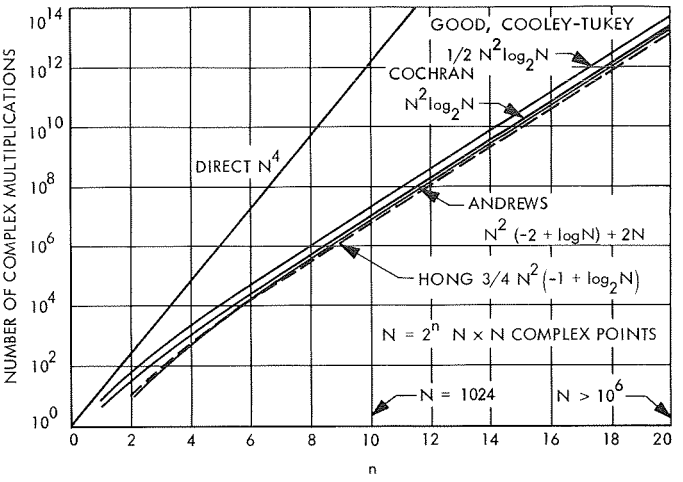


Fig. 1. Number of complex multiplications for fast two-dimensional Fourier transform

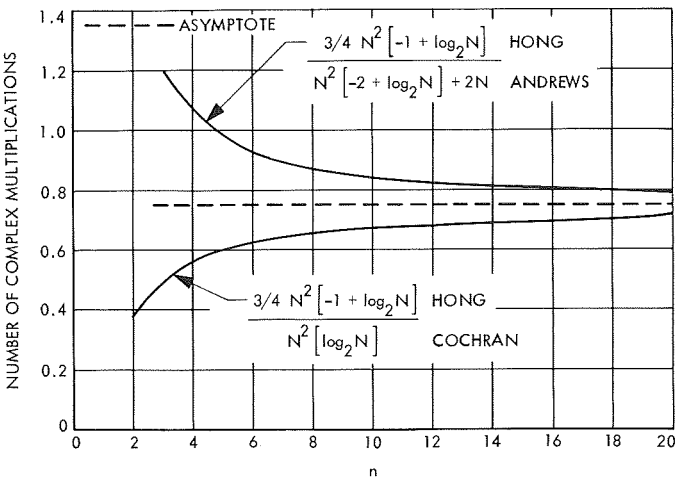


Fig. 2. Comparison of the number of complex multiplications required. Both ratios tend to 0.75

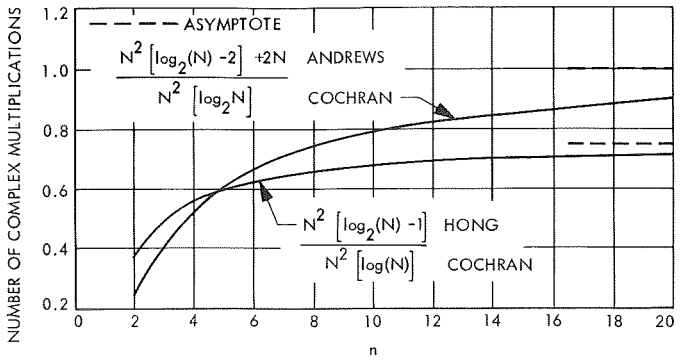


Fig. 3. Number of complex multiplication requirements compared to Cochran

Comparison with Andrews shows

$$\lim_{N\to\infty} \frac{\frac{3}{4}N^2[\log_2(N)-1]}{N^2[\log_2(N)-2]+2N} = \lim_{N\to\infty}\frac{3}{4}\frac{\log(N)-1}{\log(N)-2+2/N} = \frac{3}{4} \quad (22)$$

A computer program using the author's algorithm is being developed, and the actual time needed for execution of the fast two-dimensional Fourier transform will be determined.

It is further evident that similar savings can be found in higher dimensional Fourier transforms.

References

1. Good, I. J., *Roy. Stat. Soc.* (London), Vol. B20, p. 361; Vol. B22, p. 372.
2. Cooley, J. W., and Tukey, J. W., "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. Comput.* Vol. 19, pp. 297–301, April 1960.
3. Cooley, J. W., "Harmonic Analysis Complex Fourier Series," SHARE Program Library SDA 3425, Feb. 7, 1966.
4. Cochran, W. T., et al., "What Is the Fast Fourier Transform?", *IEEE Trans. Aud. Electroacoustics,* Vol. AU-15, No. 2, June 1967, pp. 45–55.
5. Andrews, H., "A High-Speed Algorithm for the Computer Generation of Fourier Transforms," *IEEE Trans. Comput.,* April 1968, pp. 373–375.
6. Pratt, W. K., Andrews, H., "Transform Processing and Coding of Images," USCEE Report 341, University of Southern California, March 1969.

B. Probability of Error of a Rank Code Detector in Gaussian and Nongaussian Noise,

H. D. Chadwick

1. Introduction

Rank permutation group coding (Refs. 1 and 2), or more simply, rank coding, is a coding technique for communications systems based on the properties of rank correlation and of permutation groups. At present, a study is under way of the applications of this coding technique to telemetry systems, particularly under conditions of nongaussian and nonstationary noise. Such conditions could be encountered in a lander-to-orbiter relay link due to doppler and multipath interference.

For this article, the probability of error of such a system was determined experimentally, using various rank codes under various noise conditions. The objective was to find which codes gave the best performance in terms of the lowest probability of error per bit at the highest information rate (bits per second). This data will be useful in the design of a proposed frequency modulated rank-coded system now being considered.

The first part of this article is a review of rank correlation and permutation group theory leading to the development of rank coding. The second part summarizes the experimental results.

2. Mathematical Background of Rank Coding

a. Rank vectors. A rank vector is formed from a sample of observations by replacing each observation x_i by its relative ranking in the sample. Thus, the smallest sample is replaced by 1; the second smallest by 2; and the largest (in a sample of size n) is replaced by n. The rank vector is a permutation of the digits $1, 2, \cdots, n$ and it can be shown that if the observations, x_i, are independent and identically distributed, *iid*, then all of the $n!$ permutations are equally probable, whatever the underlying probability density $f(x)$. This fact forms the basis for many nonparametric hypothesis tests.

b. Rank correlation. Rank correlation is a measure of the closeness of two rank vectors in much the same way that the linear correlation or scalar product is a measure of the closeness or distance between ordinary vectors. The rank correlation is usually measured in one of two ways: (1) the Spearman coefficient, which is equivalent to the Euclidean distance between the vectors, and (2) the Kendall coefficient, which can be shown to be very similar to the Hamming distance between two binary code word vectors. A quantity closely related to the Kendall rank correlation coefficient is called Q. The distance Q, between any rank vector permutation and the identity vector $\mathbf{1}$, which is the permutation $(1, 2, \cdots, n)$, can be considered the weight w of the permutation. This weight is closely analogous to the weight of a binary code word and can be computed in the following manner:

For every digit in the permutation, count the number of digits of higher numerical value which are to the left of this digit. The sum of the counts for all digits in the permutation is the weight w of the permutation.

Example

Consider the permutation (54213) with $n = 5$. There is one higher digit to the left of 4, two higher digits to the left of 2, three higher digits to the left of 1, and two higher digits to the left of 3. The total is 8, which is the weight of this permutation. The weight of a permutation may also be computed by the formula

$$w = \sum_{j=2}^{n} \sum_{i=1}^{j-1} m_{ij}$$

where

$$m_{ij} = \begin{cases} 1 & x_i > x_j \\ 0 & x_i < x_j \end{cases}$$

The value of w for a given permutation is the number of transpositions of two digits differing by one required to transform the permutation into the identity vector, $\mathbf{1} = (1, 2, 3, \cdots, n)$, and since the process is reversible, it is the number of such transpositions needed to transform the vector $\mathbf{1}$ into the original permutation.

c. Permutation operation. One permutation may operate on another by application of a substitution operator. The notation $AB = C$ indicates that the permutation A operates in the permutation B to form the permutation C. In general, the permutation $A = (a_1, a_2, \cdots, a_n)$ operating on the permutation $B = (b_1, b_2, \cdots, b_n)$ to form the permutation $C = (c_1, c_2, \cdots, c_n)$ means to replace the integer i where it appears in B by the integer a_i to form C; or if $b_j = i$, then $c_j = a_i$.

Example

$$(53124)\,(13542) = (51423)$$

where

$$b_1 = 1 \rightarrow c_1 = a_1 = 5$$
$$b_5 = 2 \rightarrow c_5 = a_2 = 3$$
$$b_2 = 3 \rightarrow c_2 = a_3 = 1$$
$$b_4 = 4 \rightarrow c_4 = a_4 = 2$$
$$b_3 = 5 \rightarrow c_3 = a_5 = 4$$

Each such operation A may be decomposed into a series of transpositions of digits differing by one in the permutation B, which transforms it into the permutation C. The minimum number of such transpositions necessary to perform the transposition is equal to the distance Q between B and C and may be computed by calculating the weight of the operation A.

d. Permutation groups. The set of $n!$ permutations of n digits $(1, 2, \cdots, n)$ together with the permutation operation described in the preceding section form a group. This group is generally called the symmetric group S_n on n letters. The following theorem, stated here without proof, is useful for determining the properties of subgroups of S_n as codes.

Theorem

The minimum distance between any two members of a subgroup of the symmetric group S_n is equal to the weight of the member with the lowest weight, excluding the identity element.

e. Rank vector codes. If a code is selected as subgroup of S_n with the property that the minimum distance between the elements is some value d, then, if $d - 1$ or fewer transpositions occur in the transmission of a code word, it will be possible to detect that an error has occurred. If a rank vector code word $R = (r_1, r_2, \cdots, r_n)$ is transmitted physically in the form

$$S_i = cr_i \qquad i = 1, 2, \cdots, n$$

where S_i is the value of the signal at the time instants $t = t_i = i\Delta T$; then, in the absence of noise, the receiver will form the rank vector of the received signal correctly, and no error will have occurred. With this signal form, the most probable type of error caused by additive noise will be such as to transpose two signal levels which are closest together, or in the received rank vector, to interchange two digits differing by one. If one such interchange occurs, the distance between the received rank vector and the transmitted rank vector will be one. Similarly, a larger number of interchanges, which have correspondingly lower probability, represent greater distances between the transmitted and received permutations.

f. Cyclic subgroups. One class of subgroups of S_n may be formed by taking all of the distinct powers of one of the elements: A, A^2, $A^3 \cdots$ where $A^2 = AA$ indicates that the permutation operates on itself. Such a subgroup always includes the identity element and always closes on itself. The members of the subgroup will be

$$A, A^2, A^3 \cdots, A^{r-1}, 1$$

where

$$A^r = 1 \text{ and } A^{r+s} = A^s$$

This subgroup has r members. The value of r depends on the generating element A and on the size n of the permutation. This method has been used to find subgroups of S_n for use as possible codes.

g. Decoding. The advantage of using a subgroup as a code is similar to that encountered in binary group codes: it is possible to construct a decoding table using the subgroup members and the permutations of weight 1 or less as coset leaders in a standard array.

Example

(1234)	(3142)	(2413)	(4321)	code words
(2134)	(3241)	(1423)	(4312)	possible
(1324)	(2143)	(3412)	(4231)	received
(1243)	(4132)	(2314)	(3421)	rank vectors
coset leaders				

In this example, the coset leaders have been chosen to be those permutations of weight 1, which represent a single transposition in the received word. Thus, any received word which contains an error of weight 1 or less may be decoded correctly by looking it up in the table and choosing the code word directly above it. In general, if all of the error patterns of weight e or less are used as coset leaders, it is possible to correct any error of weight e or less.

Such decoding may be performed mechanically using the following algorithm:

(1) If R, the received rank vector, is a code word, accept that code word as the decoded output.

(2) If R is not a code word, form the products RR_i^{-1} where R_i are the possible code words and calculate the weight of each product. The accepted code word is that R_i for which the weight of the product is 1 or less.

Example

In this example, the code described above is used, and it is assumed that the code word (3142) was transmitted and the permutation (2143) was received. The error of weight 1 was caused by the interchange of the digits 2 and 3. The code words and their inverses are:

$$R_1 = (1234) \qquad R_1^{-1} = (1234)$$
$$R_2 = (3142) \qquad R_2^{-1} = (2413)$$
$$R_3 = (2413) \qquad R_3^{-1} = (3142)$$
$$R_4 = (4321) \qquad R_4^{-1} = (4321)$$

To apply the algorithm, one would multiply each of the inverted code words by the received vector (2143)

$$(2143)\,(1234) = (2143) \qquad \text{weight} = 2$$
$$(2143)\,(2413) = (1324) \qquad \text{weight} = 1$$
$$(2143)\,(3142) = (4231) \qquad \text{weight} = 5$$
$$(2143)\,(4321) = (3412) \qquad \text{weight} = 4$$

Of these, the product with weight $w = 1$ or less is that using R_2^{-1}, hence $R_2 = (3142)$ is the correct code word.

3. Experimental Results

a. FM rank vector system. A communication system is presently under investigation based on the rank-coding technique combined with a frequency modulation scheme. In this system (Fig. 4), the basic rank-coded signals $s(t)$ are frequency-modulated on a carrier for transmission. The detector consists of a binary counter that counts the number of zero crossings in a time interval ΔT and then is reset to zero for the next time interval. The output of the counter, in the absence of noise, is thus proportional to the average frequency of the incoming signal over the time interval ΔT. A received code word consists of n such time intervals over a period T, or $T = n\Delta T$, where n is the sample size as used in the previous section.

As shown in Fig. 5, if $s(t)$ is an ascending ramp, corresponding to the transmitted rank vector $R_T = (1, 2, \cdots, n)$, then the received signal, without noise, will also be an ascending ramp with the received rank vector $R_R = (1, 2, \cdots, n)$. With a descending ramp signal $R_T = (n, n-1, \cdots, 1)$, the received rank vector will also be identical. In the binary case, these two signals correspond to the largest minimum distance code for any n, with the distance being given by $n(n-1)/2$. For m-ary codes, $m > 2$, the situation is more complex, but it may be shown (Ref. 3) that the minimum probability of error (when the noise in the samples is gaussian) is obtained when the average distance between observations differing by one in rank is a constant. While it has not been proved for nongaussian noise, it is assumed here to be

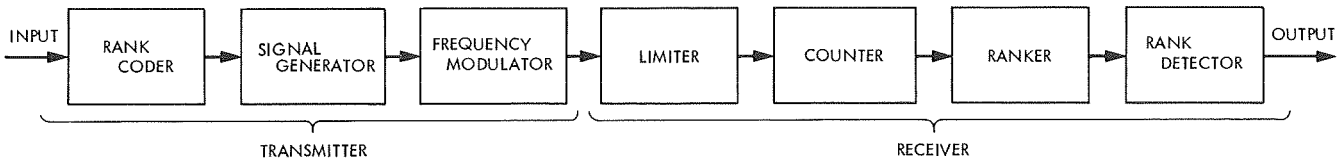


Fig. 4. FM rank-coded system block diagram

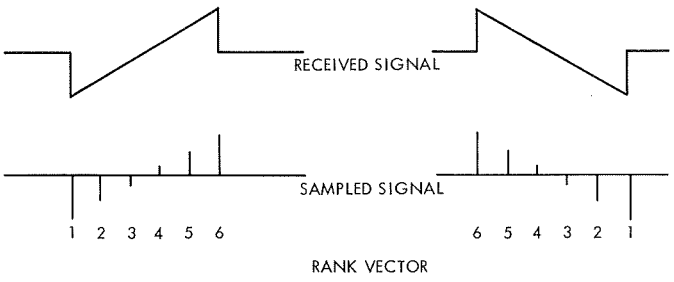


Fig. 5. Rank vector formation—binary signals

true for all noise distributions. For such a condition to be true at the output of the counter, the signal must be designed such that

$$S_k = \int_{(k-1)\Delta t}^{k\Delta t} s(t)\,dt = \ell c \text{ for } k = 1, 2, \cdots, n$$

where ℓ is an integer corresponding to the rank of the kth observation and c is a constant proportional to the modulation index of the FM signal. There are many such signals, but the two simplest are when $s(t)$ is composed of step segments and when $s(t)$ is composed of ramp segments, as in Fig. 6.

Work is presently under way to determine the statistics of the counter output for the ramp type signals in the binary case. With this information, it will be possible to predict the mean and variance of the counter output for various combinations of input noise power density, carrier power, modulation index and input filtering.

Two facts should be noted about the noise at the output of the counter. The first is that it is definitely nongaussian, being equivalent in many ways to the noise at the output of an FM filter–limiter–discriminator combination. The second is that it is sampled by the action of the counter, and therefore the mean and variance of the sample observations may be used in determining the behavior of the detector rather than the noise power density.

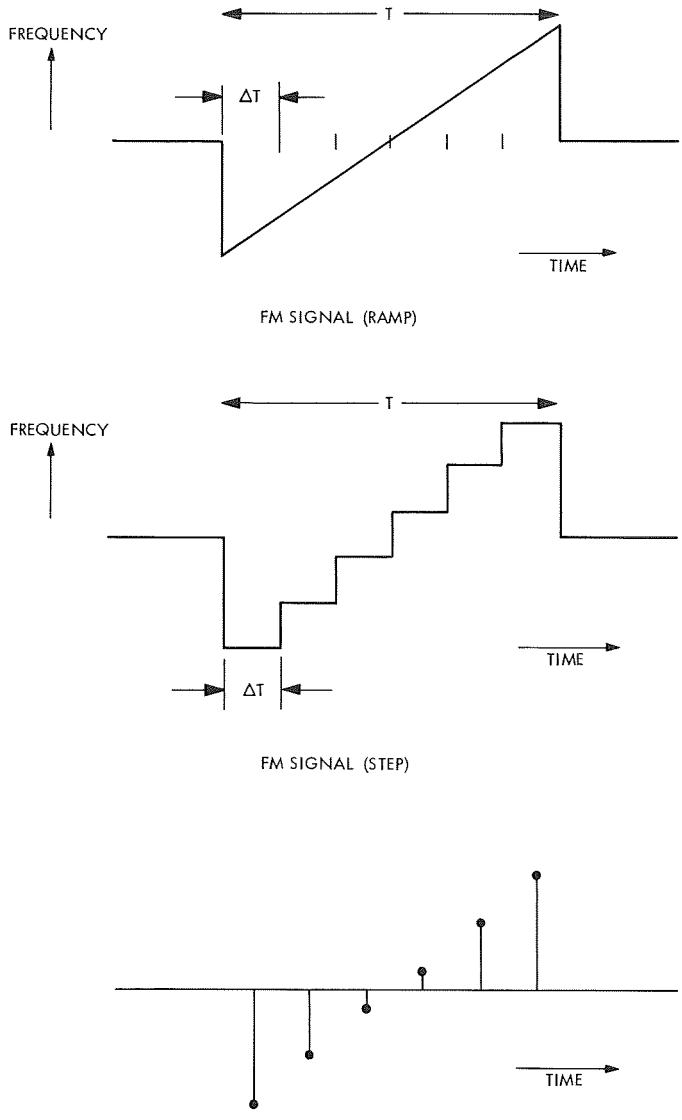


Fig. 6. Averaging and sampling by counter

b. Experimental conditions. The objective in obtaining these results was to determine the effect of the sample size n, and the number of code words in the set m on the performance of the rank-coded detector as a function of

the mean and variance of the detector input observations. These results can then be used in conjunction with the results on the statistics of the counter output to provide information on the overall performance of the receiver as a function of the input signal-to-noise ratio. This information will then be useful in the optimization of the overall receiver.

The results were obtained by the Monte-Carlo method of simulation. A large number (10,000 in most cases) of trails were run at combinations of m and n and with different noise distributions. The sample observations at the detector input (counter output) were assumed to be composed of the sum of a signal term, given by

$$S_i = br_i^T - a \qquad i = 1, 2, \cdots, n$$

where

$$R_T = \{r_i^T\}$$

is the transmitted rank vector, **a** is a constant chosen so that

$$\sum_{i=1}^{n} S_i = 0$$

and **b** is chosen so that the average power in an ascending or descending ramp signal over the time interval T is unity. The signals are assumed to be transmitted at a constant rate proportional to $1/T$ independent of either n or m so that as n is increased, the average separation between sample observations b is reduced. As m is increased, the information rate is also increased and is given in bits per second by $1/T\log_2 m$.

Four different noise distributions were tested:

(1) Gaussian, with mean zero and variance σ^2, in which case the signal-to-noise ratio per received word is defined as

$$\gamma_w^2 = \frac{1}{\sigma^2}$$

and the signal-to-noise ratio per bit at

$$\gamma_B^2 = \frac{\gamma_w^2}{\log_2 m}$$

(2) Cauchy, with pdf

$$f(x) = \frac{\lambda}{\pi(x^2 + \lambda^2)}$$

This distribution does not have a finite variance, so the parameter λ was used to measure the signal-to-noise ratio

$$\gamma_w^2 = \frac{1}{\lambda^2}$$

(3) Impulsive, with probability $p/2$ of a large positive impulse occurring at a sampling time, probability $p/2$ of an equal negative impulse, and probability $1 - p$ of no impulse or signal only. The signal-to-noise ratio was defined for this case as $\gamma_w^2 = 1/10p^2$.

(4) Uniform, with spread $-a/2$ to $a/2$, the signal-to-noise ratio defined as

$$\gamma_w^2 = \frac{1}{\sigma^2} = \frac{12}{a^2}$$

In each case, the probability of error per bit was calculated from the experimental probability of error per word and the relation

$$p_E^B = 1 - (1 - p_E^w)^{1/\log_2 m}$$

This makes it possible to compare m-ary systems for different m on a per bit basis.

c. Experimental results. The results are shown in Figs. 7 to 11. Figures 7 to 10 illustrate the effect of increasing n for a fixed signal set size m under the four different types of noise. In all cases, the probability of error per bit is decreased with an increase in n. If it is assumed that the signal interval T is held constant, as was true in this experiment, then an increase in the number of samples per interval n improves the performance of any rank-coded system at the expense of additional complexity in the decoding process. The experimental setup included the effect of increased n on the signal terms at the input (the separation between levels is decreased) but did not include the effect on the noise variance. Presumably the noise variance will also be reduced at increased n, but because the amount of change is unknown at the present this was not included as a factor.

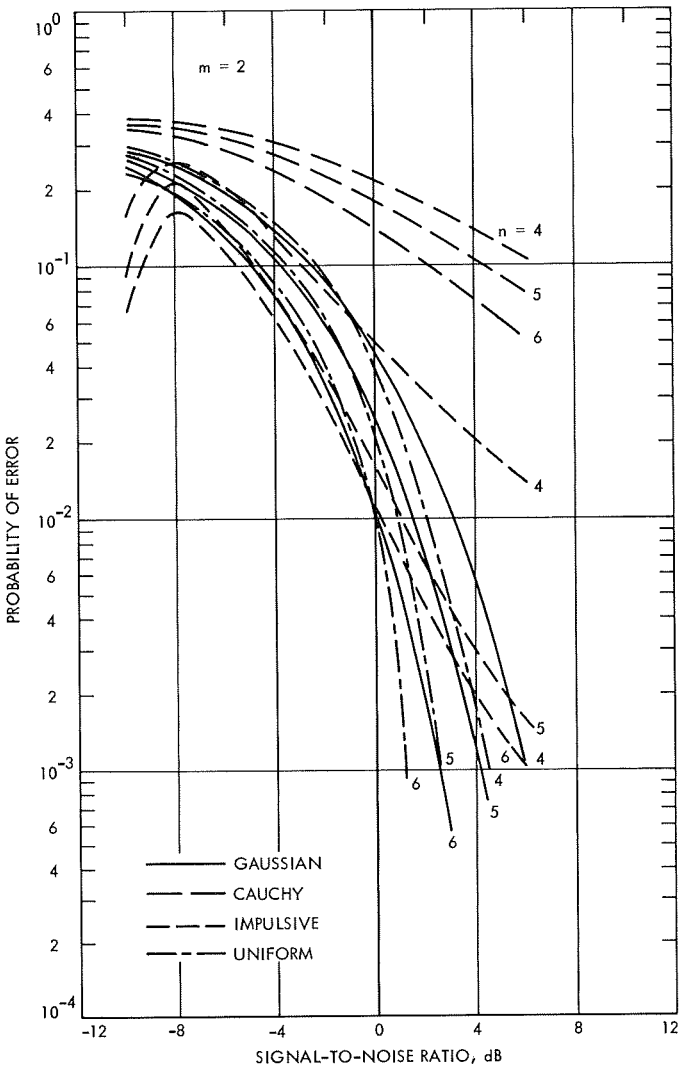


Fig. 7. Increasing n for fixed set size, $m = 2$

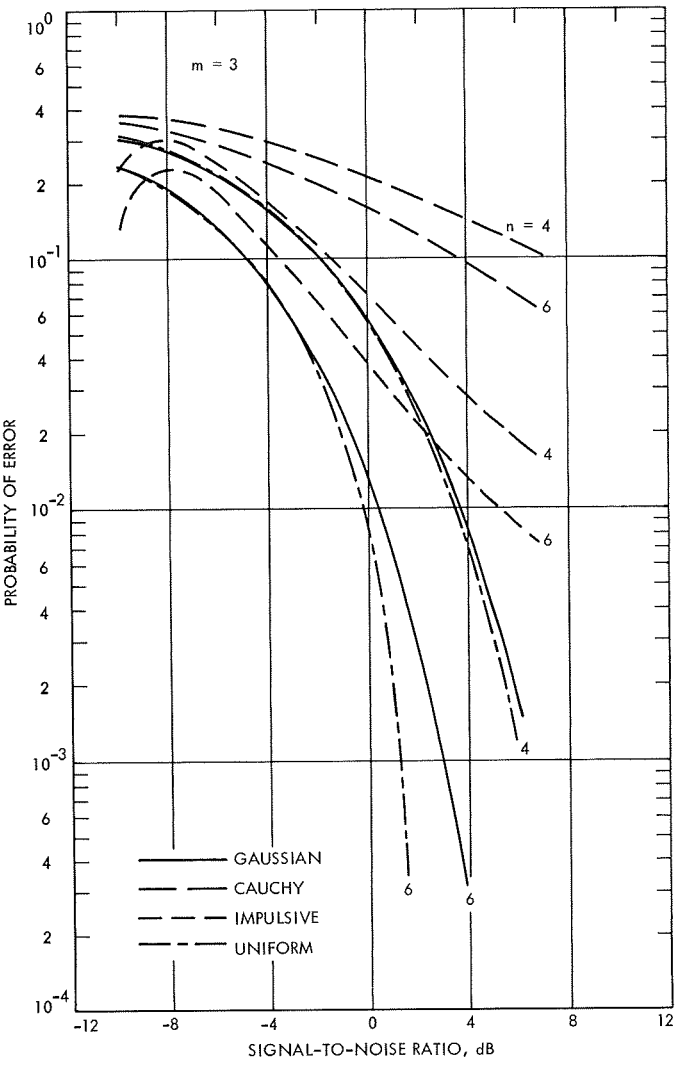


Fig. 8. Increasing n for fixed set size, $m = 3$

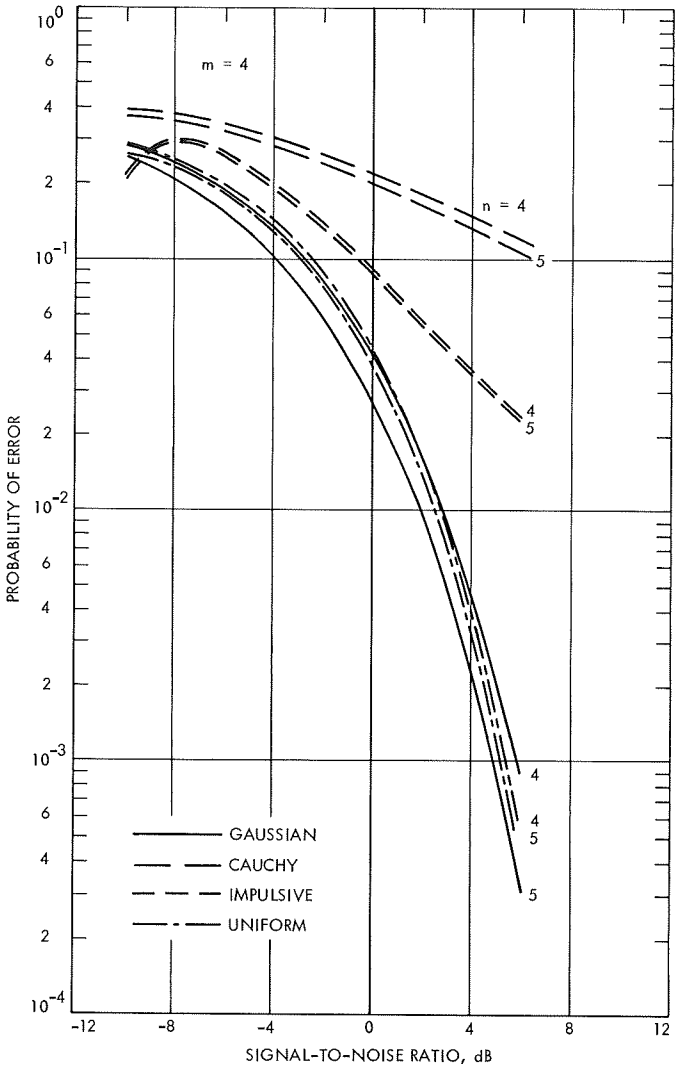


Fig. 9. Increasing *n* for fixed set size, *m* = 4

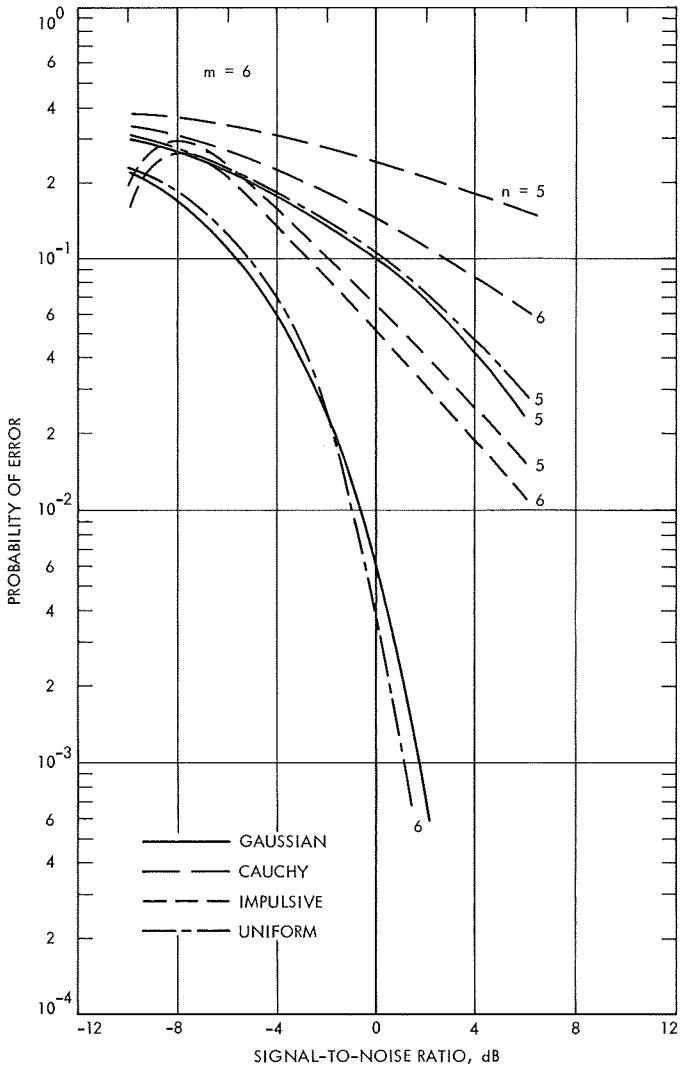


Fig. 10. Increasing *n* for fixed set size, *m* = 6

Figure 11 illustrates the performance at constant n when m is increased. These curves are plotted against the relative information rate $\log_2 m$. No simple generalization may be made about these results, but it may be seen that the code with $n = 6$, $m = 6$, offers the best probability of error of those codes tested at a relatively high information rate under gaussian and Cauchy noise[1]. The most important observation is that the probability of error does not necessarily increase with increased information rate using the rank coding scheme. The codes used here were generated by the cyclic method described in *Paragraph 2f* and not in general as "close packed" as the theoretical maximum. By finding codes with a greater m for a given n and d, it would be possible to improve the performance of the rank-coding system further in terms of information rate versus probability of error.

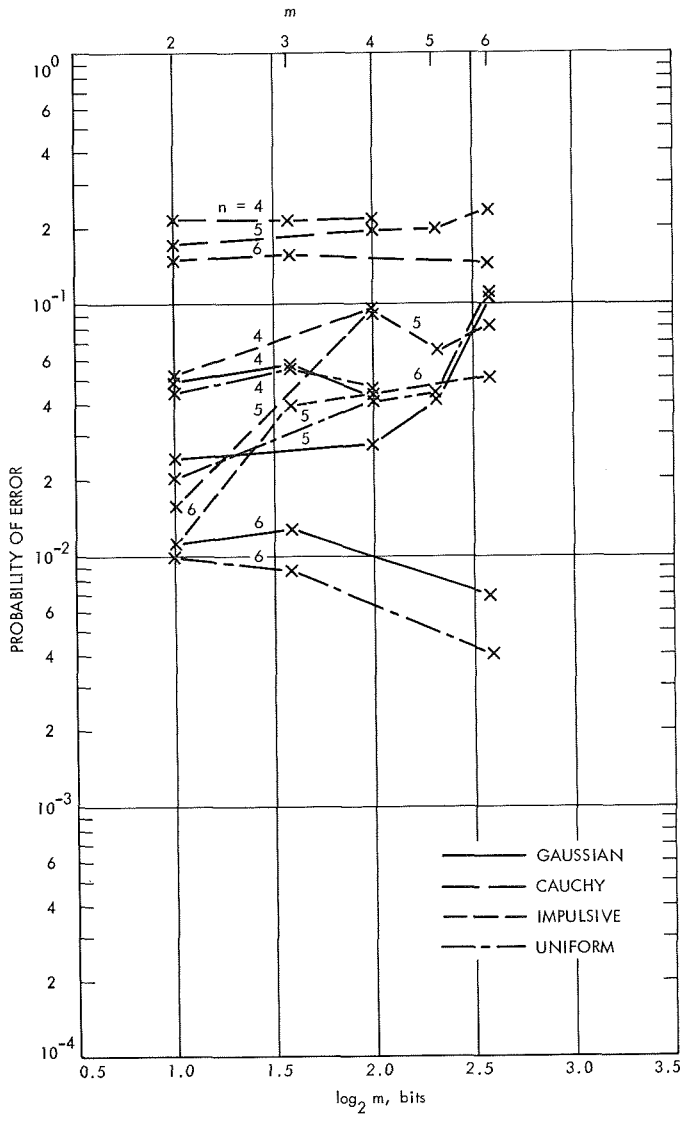


Fig. 11. Performance at constant n when m is increased

4. Conclusions

The results presented in this report provide a means for predicting the performance of the proposed FM rank-coded communications system. The results show that coding may be used to improve both the information rate and the probability of error, and that an increase in the number of samples per signal n always results in improved performance. The best code discovered of those tested was that with $n = 6$ and $m = 6$.

References

1. Chadwick, H., and Kurz, L., "Rank Permutation Group Codes Based on Kendall's Correlation Statistic," *IEEE Trans. Inform. Theory*, IT-15, pp. 306–315, Mar. 1969.

2. Chadwick, H., and Kurz, L., "Rank Permutation Group Codes Based on Kendall's Correlation Statistic," Technical Report 400-153. New York University Laboratory for Electroscience Research, Dec. 1966.

3. Stoll, E., and Kurz, L., "Suboptimum Rank Detection Procedures Using Rank Vector Codes," *IEEE Trans. Commun. Tech.*, COM-16, pp. 402–410, June 1968.

[1]While the probability of error with Cauchy noise appears high, it should be remembered that the theoretical variance of this distribution is infinite, and that, therefore, this may be looked upon as a worst-case situation.

XI. Spacecraft Power

GUIDANCE AND CONTROL DIVISION

A. Tracking and Data Relay Satellite Network,
M. Swerdling

1. Introduction

The purpose of the study described in this article was to examine several options for a Tracking and Data Relay Satellite Network (TDRSN) and to determine the feasibility of the telecommunications system. Emphasis was placed on the Tracking and Data Relay Satellite (TDRS) portion of the TDRSN. This study was undertaken at the request of NASA Headquarters and the Goddard Space Flight Center (GSFC). The study was conducted under the direction of GSFC.

The two primary objectives of the study were:

(1) To provide technical and cost data on various options for alternative TDRS configurations.

(2) To provide more detailed information on the telecommunications system for one of the alternative TDRS configurations.

2. TDRSN Concept

The concept of using a network of synchronous satellites for the purpose of relaying data from low-altitude orbiting spacecraft to the ground was conceived in 1963 by Dr. F. O. Vanbun of GSFC. Interest in a TDRSN is motivated by the limited radio contact presently available between low-altitude earth satellites and the ground. The concept of the TDRSN is illustrated in Fig. 1. This figure illustrates several aspects of the Dual-Purpose Configuration from this study. It should be noted that data generated on a user spacecraft can be transmitted first to a relay satellite, and then retransmitted to a ground station, as shown on the left half of the figure. In this way, it is possible to provide 100% radio contact (for an entire user orbit) between user satellites and the ground. Interest is also motivated by the lower predicted overall tracking and data acquisition operating costs of a TDRSN as compared with the present ground-based networks. The prospect of providing greatly improved tracking and data acquisition service, while at the same time reducing associated costs, makes the TDRSN extremely attractive. Moreover, a third benefit resulting from the implementation of a TDRSN would be a reduction in the number of overseas ground stations.

With regard to the service, it is relatively easy for the TDRSN to provide service to low altitude users, and it becomes more difficult for a TDRSN to serve higher altitude users. On the other hand, three ground stations can

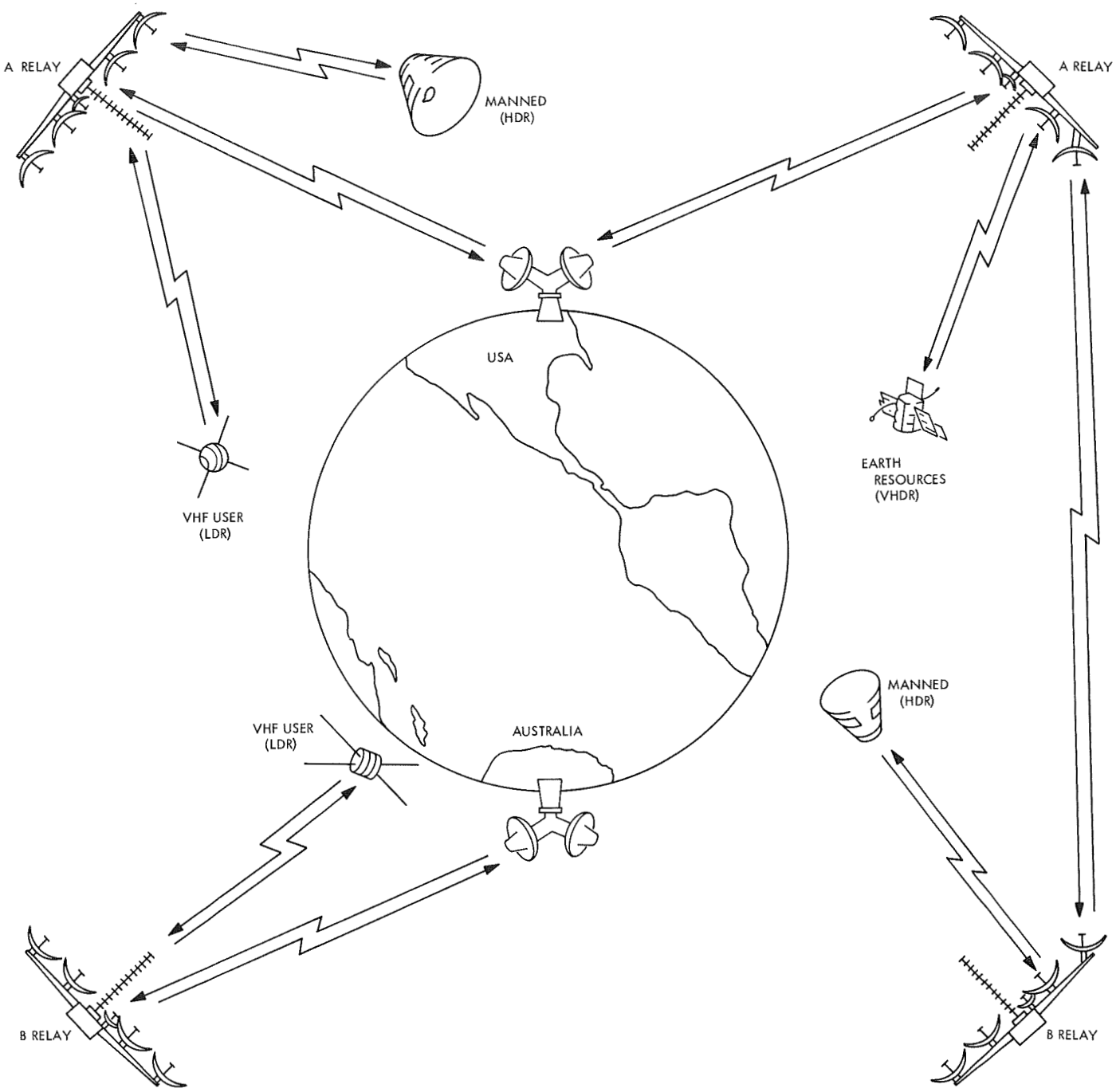


Fig. 1. TDRSN concept

provide 100% coverage for high altitude users, and it becomes more difficult for a ground network to service lower altitude users. Thus, there may be an optimum altitude which should be served by a TDRSN. Users below this dividing altitude would be served by the TDRSN, and users above this dividing altitude would be served by a ground network. For this study, the dividing altitude was taken to be 3,000 km.

Many different users must be serviced by the TDRSN, but from a design point-of-view, all users can be classified as one of two types. The first type is the low data rate (LDR) user, which is characterized by transmitting and receiving information through an omni-directional antenna. The second type is the high data rate (HDR) user, which is characterized by transmitting and receiving information through an oriented, narrow-beamwidth

antenna. In this study, a third type of user was identified as a very high data rate (VHDR) user; however, in reality, this type is the same as the HDR user except that greater communication bandwidth is required.

Four alternative configurations were investigated in sufficient depth to provide comparative technical and cost information for various levels of capability.

The term Dual-Access Configuration indicates that a single TDRS can simultaneously serve any combination of two LDR or HDR users. This configuration was designed to serve manned missions, and was included in the study in order to assess the cost of satisfying only the manned requirements. This concept is illustrated in Fig. 2.

The Subsynchronous Configuration employs TDRS located in 8-h, subsynchronous orbits. There existed a question as to whether or not a subsynchronous configuration could provide acceptable service at a lower cost than a TDRSN utilizing synchronous TDRS; the study showed that the Subsynchronous Configuration is unsatisfactory.

The Inter-Relay Configuration contains an "inter-relay link" by which data received at one TDRS from a user spacecraft is relayed through a second TDRS prior to being received at the ground station. The use of this inter-relay link makes it possible to provide 100% coverage without an overseas ground station. The Inter-Relay Configuration is illustrated in Fig. 3.

The Dual-Purpose Configuration contains a "dual-purpose link" that can be used either as an inter-relay link or as another HDR user link. This configuration was designed to provide maximum flexibility in operation since future requirements for the TDRSN are uncertain at the present time. This configuration is illustrated in Fig. 4.

A summary of the characteristics of the three feasible configurations is given in Table 1. The weight of the Inter-Relay Configuration is such that an *Atlas/Centaur* launch vehicle could be utilized if a 28.5-deg inclined relay orbit is used instead of an equatorial orbit. The weight of the Dual-Purpose Configuration is such that

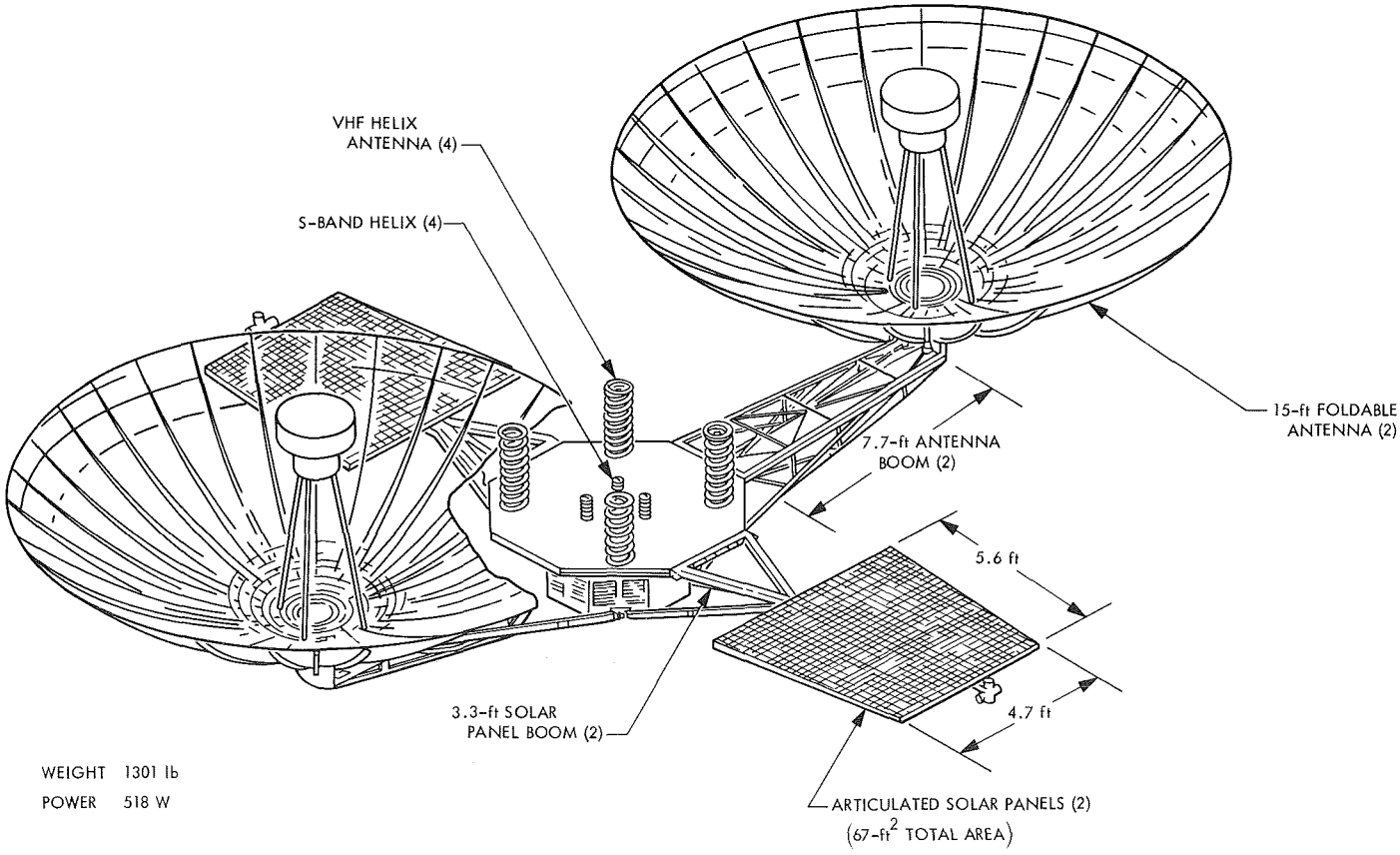


Fig. 2. Dual-Access Configuration

the *Titan IIIC* provides a large weight margin for a 28.5-deg inclined orbit, but is inadequate for an equatorial orbit. All of the configurations were designed to provide 100% coverage, and the power systems were sized to provide uninterrupted, complete service during periods of sun occultation. All three configurations require two 85-ft-diam antennas (without spares) at each ground station.

3. Guidance and Control

a. Dual-purpose power subsystem. The power subsystem described here is considered a worst-case design of the alternate configurations. The preliminary design was based on a constraint to provide continuous power throughout a 24-h synchronous orbit for a 5-yr operating period while performing at a 28.5-deg equatorial inclination. The battery was to provide a total of about 763 discharge/charge cycles throughout the mission.

Sizing of the power subsystem resulted from a survey (Table 2) of using subsystems. Table 2 is based on the largest power requirement, which occurs when a TDRS B relay is operating in an inter-relay mode with a TDRS A relay (Fig. 1). The total communication requirement, including conversion and distribution losses, represents 79% of the total TDRS power requirement (including battery charging power). Conversion efficiencies were assumed to be 90% for those loads requiring close regulation, i.e., within ±1–2% range. Distribution losses (based on actual flight programs) were assumed to be 3%.

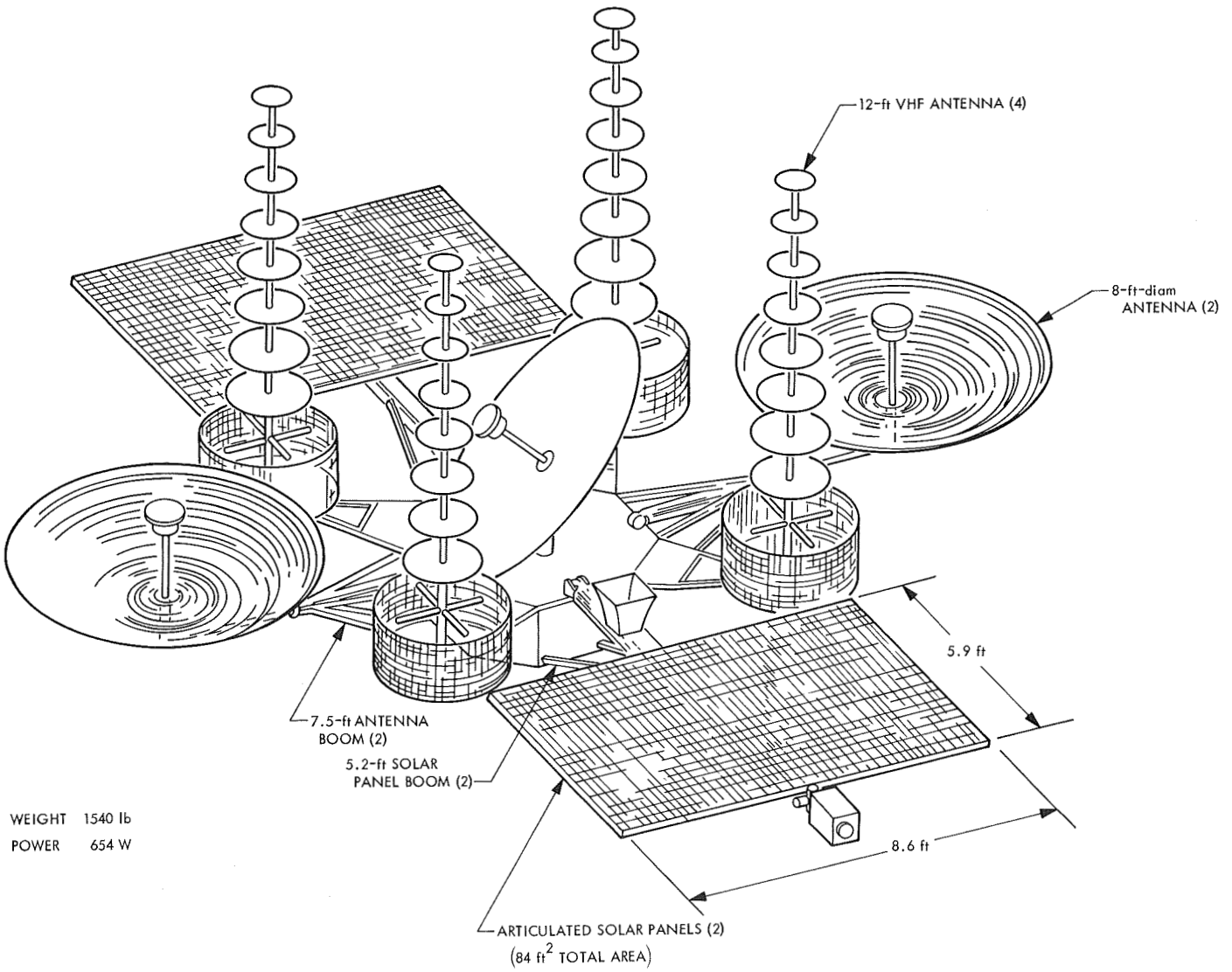


Fig. 3. Inter-Relay Configuration

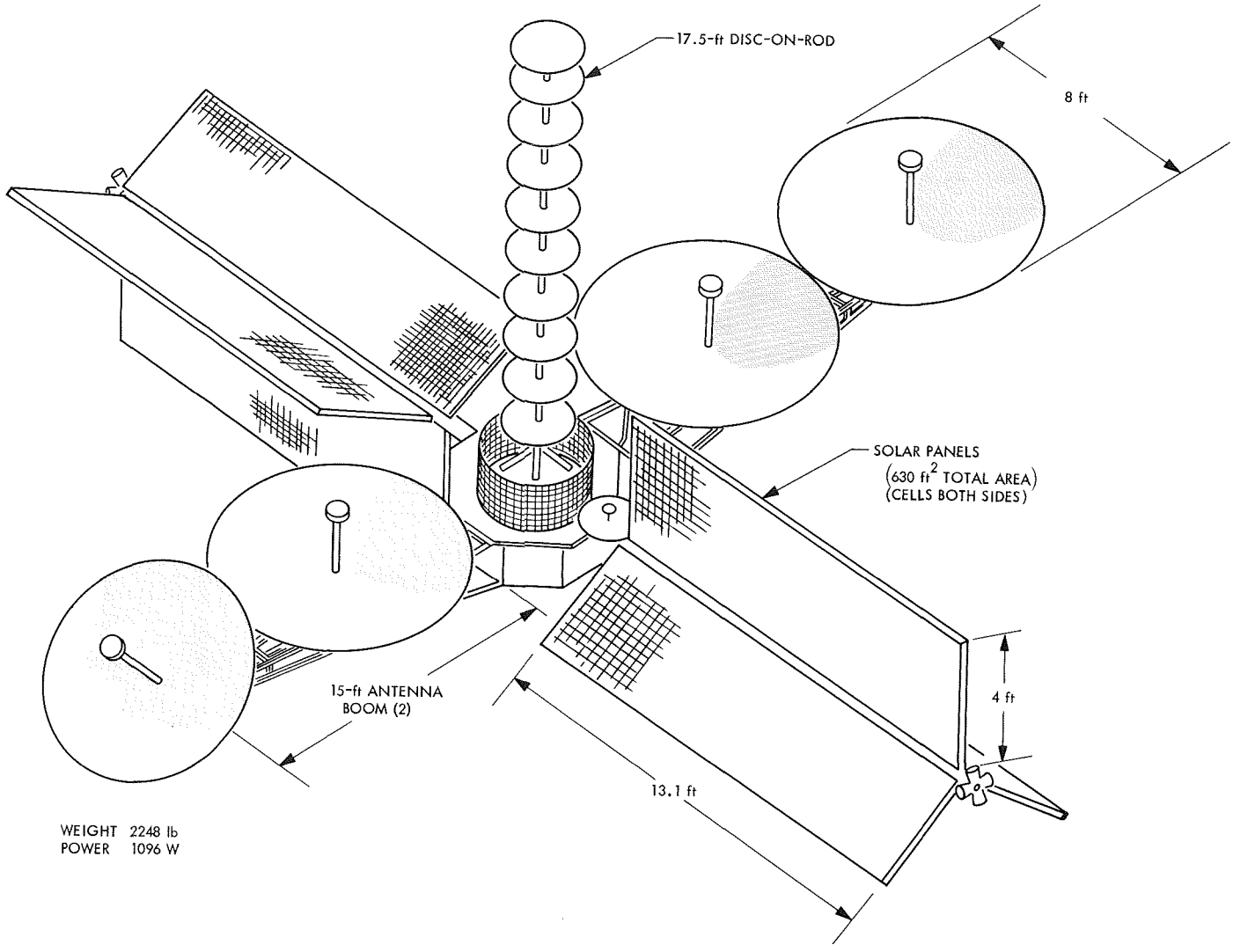


Fig. 4. Dual-Purpose Configuration

b. Solar panels. Fixed lightweight solar panels were selected for the Dual-Purpose Configuration. This selection was based, primarily, on the state-of-the-art of rotating panels, the ability of which to survive a 5-yr mission is questionable. The penalty for this decision is approximately 154 lb for each TDRS. However, if demonstration of the rotary power transformer (RPT) proves successful prior to the technology cutoff date (March 1971), a change to the rotating panel and RPT is deemed advisable.

Active solar-panel area was determined by considering the end of the 5-yr mission 1096-W power requirement (Table 2) in addition to an assumed 20% degradation in performance due to radiation, uncertainties, etc. Increasing the panel area due to operating with an initial ecliptic inclination of 20 deg and decreasing to 11 deg after 5 yr was not required, since the peak demands for battery charging, which occur at alternate quarter-year intervals relative to the peak solar-panel inclinations, were greater than the losses caused by the inclined solar panels.

The solar-panel power output is estimated to be 1370 W (66°C) initially, degrading to 1096 W at the end of 5 yr. Performance is based on the use of a lightweight solar panel consisting of an 8-mil N/P silicon solar cell and a 6-mil coverglass filter. The total active cell area required is estimated to be 163 ft^2, based on the area exposed to the sun of the selected 6-panel configuration. However, because the panels are of the fixed type on a 3-axis stabilized TDRS, total active solar cell area is increased to

Table 1. Comparison of alternative configurations

Configuration	Weight, lb	Power, W	Attitude control	TDRS orbit	Number of TDRSs	Number of ground stations	Antennas	Launch date (start in July 1970), 1974
Dual access	1301	518 (rotating solar panels, NiCd batteries)	3-axis stabilized with momentum wheels	Synchronous (28.5-deg inclined)	4 (90-deg apart)	2 (U.S. and Australia)	Two 15-ft unfurlable dishes with dual UHF/X-band feed 4-element helix array at UHF 4-element helix array at S-band	March (*Atlas/Centaur*)
Inter-relay	1540	654 (rotating solar panels, NiCd batteries)	3-axis stabilized with momentum wheels	Synchronous (equatorial)	4 (asymmetrically placed)	1 (U.S. only)	Three 8-ft rigid dishes at X-band 4-element disc-on-rod array at VHF Horn at X-band	December (*Titan IIIC*)
Dual purpose	2248	1096 (fixed solar panels, NiCd batteries)	3-axis stabilized with momentum wheels	Synchronous (28.5-deg inclined)	4 (90-deg apart)	2 (U.S. and Australia)	Four 8-ft rigid dishes at X-band 1-element disc-on-rod at VHF Dish at X-band	December (*Titan IIIC*)

Table 2. Dual-purpose power requirements

System	Power		Voltage, Vdc	Peak duration
	Continuous, W	Peak, W		
Central computer and sequencer	35	39	28 ±2%	During launch only
Antenna pointing	23.0	23	28 ±4V	
Propulsion	0	40	28 ±2%	180 s per use
Pyro	0	2	28 ±2%	
Communications	198	198	28 ±1%	
Communications	625	625	28 ±4V	
Attitude control				
Electrical wheel drive	4.5	12	28 ±1%	
Inertia wheels	1	23	28 ±1%	
Horizon scanner	3	6	28 ±1%	
Gyro compass	2	4	28 ±1%	
Valves (6 on at once)	0.5	20	28 ±4V	20 ms
Power conditioning and distribution losses	61	75		
Battery charging (c/10)	143	143		
Total	1096			

565 ft². Assuming a packing factor of 90%, the gross solar-panel area is 630 ft², divided into 6 panels with solar cells on both sides. The solar panel configuration is shown in Fig. 4. The total solar-panel weight, including cells, filters, adhesives, substrate, etc., is estimated to be 284 lb, based on the JPL/Boeing-developed, lightweight solar-panel technology. The overall specific power rating of this concept is 1.75 W/ft² (end of 5-yr mission). The specific power density rating is 3.86 W/lb.

Fixed solar array considerations. The use of solar panels in the fixed array previously described requires that the optimum location and orientation of the solar panels be determined. Also, the effects of shadowing must be considered. The form of the array shown in Fig. 5 was selected after considering various alternatives. A preliminary performance analysis of this solar array is given in the following material. The array consists of six 48 × 158-in. double-sided solar panels arranged in two sets of three panels each, with the two sets being rotated 60 deg relative to each other about the central axis. Panel absorptivities and emissivities were assumed to be 0.8, a value consistent with previous array designs. Power output of a flat array with normal solar incidence, a temperature of 60°C, and including a 20% degradation allowance, is 8.0-W/ft². Other configurations or orientations will cause changes in power capability by altering the effective intensity and/or temperature. Power capability was assumed to be directly proportional to the effective solar intensity (cosine law with incidence angle) and to have a temperature coefficient of −0.56%/°C, a value consistent with 8-mil 2 × 2 cm cells.

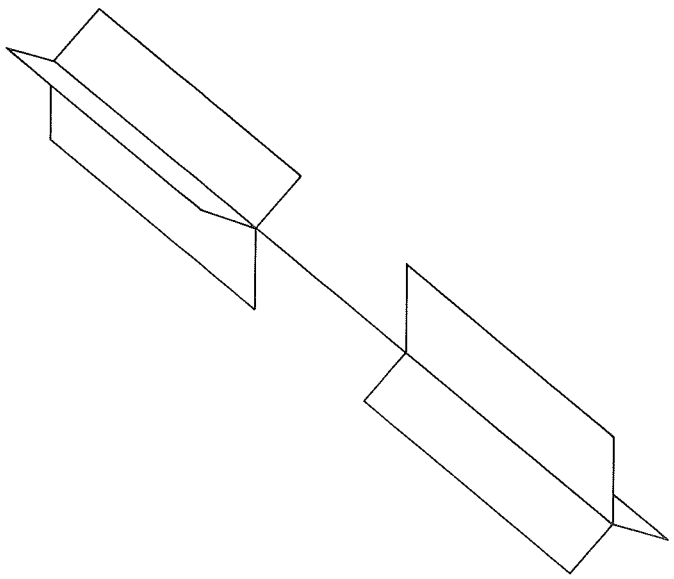


Fig. 5. Fixed solar-panel array configuration

Two orientations were considered for the case where the solar flux is normal to the array axis. The first of these is shown in Fig. 6a. Figure 6b is equivalent if edge heating of the panel nearest the sun is neglected. Using approximate expressions for the heat transfer between panels, the temperature of cells near the axis and centrally located axially on the two illuminated panels was found to be about 70°C. This temperature was used to calculate power capability, which, when including the current reduction caused by the 30-deg off-normal solar flux incidence, resulted in a power density of 6.53 W/ft². The power capability of the array for this orientation is 1370 W. Cells near the outside edges of the panels would operate at temperatures several degrees cooler than those near the center of the array; however, the corresponding increase in power capability (on the order of 1%) has been neglected.

The second orientation is shown in Fig. 7. In this orientation, one panel receives the full solar flux intensity and a second receives only half the normal flux intensity. Cell temperatures near the axis of the first panel were calculated to be 68°C; those on the second panel were 34°C. The corresponding power densities for the two panels, considered separately, were 7.65 and 4.57 W/ft². Unfortunately, the maximum power points of the two panels do not occur at the same voltage, and these two numbers cannot be combined to determine the total power capability if the panels are paralleled electrically. Assuming that the cooler panel is operating at essentially its short-circuit current, its power capability at the maximum power voltage of the warmer panel is approximately 4.2 W/ft². Using this figure, the total power capability of the array in this orientation is 1238 W.

Shadowing effects on the array. As the array rotates from the first orientation to the second during the TDRS

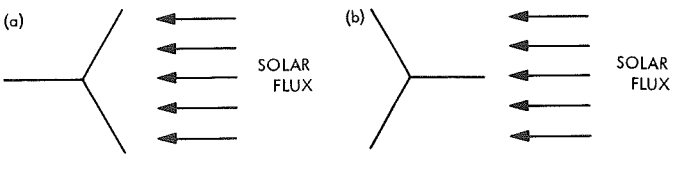


Fig. 6. Orientation 1 for fixed solar-panel array

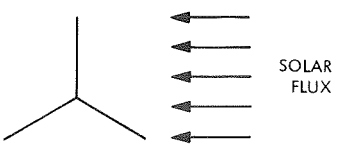


Fig. 7. Orientation 2 for fixed solar-panel array

orbit, the panel whose edge is nearest the sun will partially shadow one of the other panels, reducing the power output of the latter. A compensating effect is the increase in power output of the panel nearest the sun. It is not clear whether or not there is a minimum power output orientation that lies between the two orientations chosen for analysis; the analysis becomes quite complicated for intermediate orientations and considerably more study would be required to resolve this point. It seems unlikely that such a minimum, if it exists, would be very deep, especially since the shadowing could occur on only one-half of the solar array.

Consideration was given to the effects of shadowing of the solar panels by the VHF antenna, which occurs during portions of the daily cycle for inclined orbits. The model analyzed is shown in Fig. 8, for which the maximum orbit inclination of 25 deg was assumed. The power loss caused by shadowing varies slightly, depending upon which half of the array is shadowed. A peak loss of 6.4% occurs if the shadow falls on the array half in which one panel extends in the same direction as the antenna; the peak loss if the shadow falls on the other array half is 4.6%. The duration of the shadowing in the former case is 1.27 h; in the latter case it is 1.5 h. In both cases, it is assumed that the antenna shadow is smaller in width than the diameter of the hub of the solar array, so that the shadow cannot fall on both sides of an array half simultaneously. Also, it is assumed that if the shadow falls on any portion of the panel as the TDRS rotates, the complete swept portion of the panel contributes no power. This is a worst-case assumption, and the power loss could probably be reduced somewhat by careful cell interconnection. It should be pointed out that no shadowing by the antenna occurs for inclinations of less than 17 deg, and that the percentage power loss is approximately a linear function of inclination angle for the range from 17 to 25 deg.

c. Batteries. Hermetically sealed nickel–cadmium (NiCd) batteries were selected for the Dual-Purpose Configuration. Sizing was based on supplying energy during a worst-case, 72-min sun occultation period. In the orbits of interest for the 5-yr mission, a total of approximately 763 varying sun occultation periods will occur. This number of discharge/charge cycles imposed on the battery is well within the present NiCd battery capability, provided temperatures and depths of discharge are favorable.

With complete system operation occurring through sun occultation periods, a worst-case energy requirement of 1140 W-h (47.5 A-h at 24 V) is demanded. This energy is

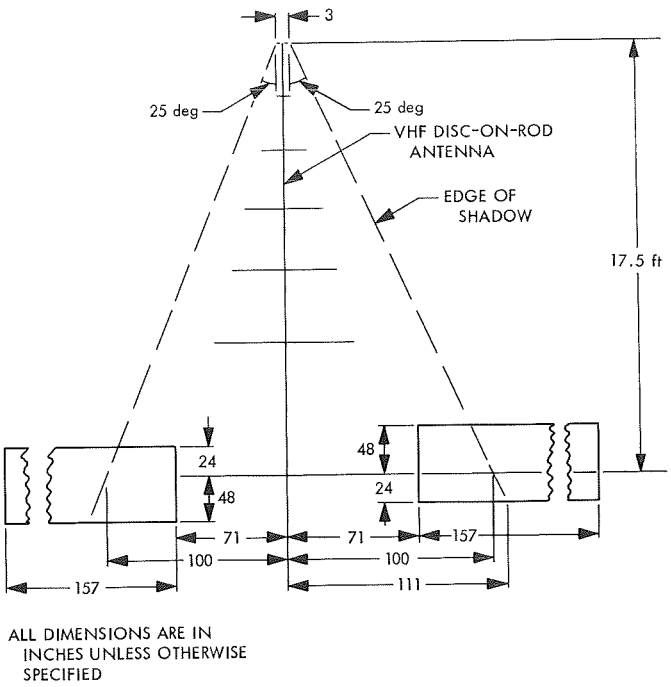


Fig. 8. Shadow model

provided by three 22-cell, 40-A-h, sequentially-recharged batteries, each discharging to an approximate 40% depth of discharge. For a c/10 charge rate, this approach reduces total battery charging power and fixed solar-panel area by 71 W and 40.5 ft^2, respectively, as compared to a two 60-A-h battery arrangement. Utilization of the 60-A-h, two-battery approach would result in an overall weight savings of 21.5 lb; however, in view of the fact that solar-panel area is more critical than the 21.5 lb saving, the three 40-A-h batteries were selected.

System operation is based on a minimum bus voltage of 24 V. Total battery terminal voltage was assumed to be 25.5 V, based on diode losses of 0.75 V and an overall system voltage drop of 0.75 V. On this basis, and combined with an average cell discharge voltage of 1.15 V/cell, the total number of cells was established at 22. Total battery weight for three banks, including an assumed 20% case weight, is estimated to be 315 lb; total volume is estimated to be 2.75 ft^3. The effective specific energy density rating is 3.62 W-h/lb.

Battery charging is accomplished at the c/10 rate, requiring 143 W of raw power (including a 90% charger efficiency). Sequential charging of the three batteries is ultilized because of the $>$ 22 h of solar energy available,

as well as a realization of a significant reduction in solar-panel area (246 ft²). At the c/10 charge rate, combined with a 40% depth of discharge, each battery will be charged in approximate 6.5 h (70°F).

d. Power conditioning and distribution. A block diagram of the dual-purpose TDRS power subsystem is shown in Fig. 9. A 28-Vdc primary bus with a tolerance of ±4 V was selected since only approximately 25% of the total loads required close regulation at 28 Vdc (±1–2%). In addition, the unregulated voltage requirement of the communication subsystem for downstream conversion to special voltages eliminated the series buildup of inefficiencies. A zener diode clamp is used to protect the bus voltage at the 32-V level when the solar-panel voltage exceeds this level. With a 28-Vdc (±1%) close-regulation requirement, the buck/boost regulator approach appears to be the best choice for regulating the battery or array voltages, if warranted.

The power source logic unit contains the dc bus from which all solar panels and battery output are connected. This unit also contains a transfer switch for selection of internal or external power and isolation diodes for solar-panel sections. The power control contains the failure sense circuit to detect out-of-tolerance conditions in the buck/boost regulator and the power control relay to switch to the standby buck/boost regulator. The power distribution unit receives signals from the Flight Command System and the Central Computer and Sequencer for distribution of power. The signals are then processed to actuate magnetic latching relays for power distribution.

e. Power subsystem weight. A weight breakdown of the dual-purpose TDRS power subsystem is summarized in Table 3. The total weight of the lightweight solar panels includes cells, filters, adhesives, substrates, and miscellaneous structures for a gross area of 630 ft² with cells on both sides of the six panels. The specific power density of the total power subsystem is approximately 1.7 W/lb (end of 5-yr mission).

f. Power subsystem tradeoffs. The following is a description of the tradeoffs performed for the selection of the primary and secondary power sources. Also discussed are tradeoffs between fixed and rotating solar panels. The tradeoffs are based on a constraint to provide continuous power throughout a 24-h orbit for a 5-yr operating period. As a guideline, a technology cutoff date of March 1971 with a launch date in December 1974 was considered for the TDRSN program.

The primary power source provides power during sunlight. The secondary power source provides power during sun occulation, and also during peak power demands.

Alternate primary power source. Long-term unmanned satellite missions, coupled with average power requirements in the TDRSN 1.0–1.4 kW range, require that some consideration be given to static primary power sources other than that of the photovoltaic variety. Static energy conversion sources are preferred to dynamic machinery for missions extending over 5 yr and up because of the increase in overall subsystem reliability. In addition, a source is required for the prime mover to drive the machine. A candidate static power source for missions of this type would be the radiosotope thermoelectric generator (RTG).

The advantages of the RTG pertaining to this mission would be the absence of solar-panel orientation requirements for rotating panels. (Fixed panels would require a 300–400% overall increase in array area over rotating panels.) Additional advantages exhibited by the RTG are the lack of sensitivity to space environmental radiation during periods of solar flare activity and environmental temperature effects. Since an RTG provides power continuously, independent of the relative position of the sun, the role of the secondary power source is significantly reduced; a battery sized for use with an RTG would only be used as a peak-load device. In addition, the effects of battery failure would be deemphasized as compared to the complete battery operation required during sun occultation for a solar-panel/battery system.

The disadvantages associated with the use of an RTG for this mission are weight, cost, and safety hazards during the prelaunch period. In terms of weight, since the power density of an RTG is approximately 1.1–1.2 W/lb, as compared to the 8.4-W/lb power density of lightweight solar panels (end of 5-yr mission), the total RTG weight would be in the range of 800–900 lb. In addition, shielding requirements and/or booms may increase these totals.

Alternate secondary power source. NiCd batteries were selected as the secondary power source for all three TDRS configurations. AgZn and AgCd batteries show higher specific energy densities than the hermetically sealed NiCd type, but the ability to survive 5 yr in space restricts the selection to the NiCd battery. NiCd batteries, considered as the "workhorse" of space batteries, have been functioning with success for several years at depths of discharges ranging from approximately 10 to 25%.

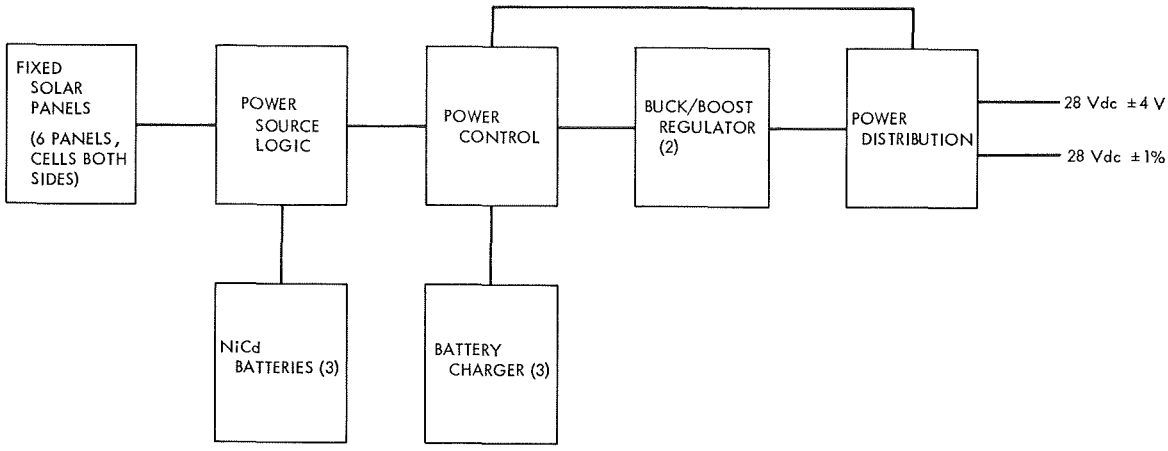


Fig. 9. TDRS Dual-Purpose Configuration power subsystem

Table 3. Dual-purpose power subsystem weight

Component	Weight, lb
Solar panels (630 ft² area)	284
3 NiCd batteries (40 A-h)	315
3 battery chargers	18
2 regulators	12
Power source logic	7
Power controller	3
Power distribution	3
Total	642

Regenerative fuel cells (RFC), considered as an alternate secondary power source, combined with solar cells provide an alternate approach that can reduce overall subsystem weight as compared to the conventional solar-cell/NiCd battery combination. The RFC alternate approach is discussed below.

The hydrogen–oxygen RFC is basically a combination of a hydrogen–oxygen primary fuel cell and a water electrolysis cell in one compact package. (These cells may be packaged separately for convenience.) Solar panels are normally used to provide power for regeneration by electrolysis of the fuel-cell-produced water into the original hydrogen and oxygen reactants. A feature of the RFC is that an approximate 100% depth of discharge could be utilized as compared to the NiCd battery, which normally would be discharged only to depths of approximately 25–40% of its capacity.

The advantages of the RFC approach are lower weight and, possibly, higher reliability when considering a 5-yr mission life. In terms of weight, the effective specific energy density rating of the dual-purpose TDRS NiCd battery system is approximately 3.62 W-h/lb. The present development of the RFC has provided sufficient evidence of achieving an effective specific energy rating of approximately 20 W-h/lb, which represents a weight improvement factor of 4 to 5 over the batteries used for the Dual-Purpose Configuration. Two RFCs (70 lb) represent a weight savings of 260 lb in lieu of the three NiCd battery packs for the Dual-Purpose Configuration. It is anticipated that if present funding by the Air Force, NASA and Comsat continues, specific energy density ratings of 30–35 W-h/lb may be achieved by March 1971 (the technology cutoff date for the TDRSN Program).

In addition to the weight advantage of the RFC for missions of this type (synchronous orbit), the effect on the primary source for recharging must also be given consideration. The RFC lends itself to synchronous orbit operation and may take advantage of the 22–23 h of sunlight available for electrolysis. To completely recharge the

NiCd battery, reasonable charge rates would be in the range of c/9 to c/10 at 70°F. In terms of charging power, considering the dual-purpose TDRS, the NiCd system would require approximately 143 W [4.0 A (32.5 V) ÷ 0.9] for a 22-cell battery at the c/10 rate. The RFC would require approximately 116 W [4.62 A (25 V)] with a 32-cell unit for approximately 22 h. The advantages and disadvantages of the RFC are as follows:

Advantages	Disadvantages
Long potential cycle life at full depth of discharge	Low watt–hour per cubic foot ratio
Very high charge/discharge rate capability	Poor ambient temperature operation capability
High watt–hour per pound ratio	Low charge retention
Good voltage regulation	
High overload capability	High pressure operation
High shock stability potential	
High temperature operation potential	
No problem of series cell balance	
Indefinite shelf life (discharged condition)	

With the established March 1971 technology cutoff date, it appears that the RFC may not be flight qualified for a 1974 launch, and, on this basis, is not recommended for the initial launch. However, the RFC should be given serious consideration for later launch dates, provided an adequate funding level is maintained for development and flight qualification.

Fixed versus rotating solar panels. The major tradeoff concerned with this decision is the additional weight required to increase the reliability for the 5-yr mission. To use fixed solar panels rather than rotating solar panels introduces a weight penalty with either conventional or the JPL newly-developed lightweight solar-panel technology regardless of whether an RPT or slip rings are used.

Considering the dual-purpose TDRS fixed solar panel total area requirement of 630 ft² with conventional *Mariner* Mars 1969 solar-panel technology, the total weight (including 18-mil N/P cells, 20-mil coverglass, and structure) is estimated to be approximately 635 lb (not including booms). The weight of the rotating panel with a single degree of freedom is estimated to be 195 lb based on a total area of 152 ft². The conventional mechanisms are estimated to be an additional 30 lb. However, the difference in weight of 410 lb (635 − 225 lb) may be reduced with the use of the lightweight solar-panel technology that includes an 8-mil N/P cell with a 6-mil coverglass filter. The dual-purpose fixed panels weigh 284 lb using this lightweight technology. Fixed and rotating solar-panel weight differences would then be approximately 154 lb, including rotary devices, or a 54% weight savings with the rotating solar panel.

The reliable operation of rotating solar panels utilizing the conventional slip rings, gears, brushless dc motors, etc., has not been fully demonstrated in space for 5-yr missions, particularly with a requirement of 1 rev/day. Since continued operation of a panel drive system in a vacuum would encounter difficulties in lubrication, it became apparent that, to ensure a long mission operation, the solar panel orientation system must be free of gears, bearings, slip rings, and sealed chambers (from which mechanical power must be transmitted). To achieve this objective, GSFC is currently developing an RPT that couples power from the rotating solar panel to the spacecraft and thereby eliminates the need for slip rings, etc. On the solar panel side, the dc output is inverted to ac for input to the RPT. The output of the RPT (square wave) on the spacecraft side may or may not be rectified to dc depending on system requirements. Current development is such that by the end of 1969, or sooner, a 500-W model will be available to demonstrate the feasibility of this approach. Using lightweight rotating panels with an RPT (130 lb) in lieu of lightweight fixed solar panels (284 lb) represents a weight savings of 154 lb (including RPT) for the Dual-Purpose Configuration. For the purpose of this study, in order to cover both types of systems, rotating solar panels were selected for the Dual-Access and Inter-Relay Configurations, and fixed solar panels were selected for the Dual-Purpose Configuration. The latter selection was made in order to consider fixed solar panels for the largest power requirement, since this appears to be the most difficult.

4. Conclusion

Fixed lightweight solar panels, in combination with hermetically sealed NiCd batteries have been selected for this mission (launch December 1974) in consideration of

the March 1971 technology cutoff date. The RTGs, although more reliable for long missions of this type, are considered too costly and heavy. For later launches, or when technology development permits, articulated solar panels and RFCs may be used with the realization of a significant reduction in weight as well as costs.

B. Heat-Sterilizable Battery Development, *R. Lutwack and W. von Hartman*

1. JPL Contract With Texas Instruments Inc.

This program to develop sterilizable nickel–cadmium cells is divided into two main sections: one for sterilizable cells and the other for sterilizable, high-impact resistant cells. In both sections, prismatic and cylindrical designs are being considered.

Additional cycle performance data have been obtained for 17- and 18-plate, 4-A-h prismatic cells, which are part of a factorial experiment designed to study the effect of sterilization. The 17-plate cells have undergone 187 cycles after sterilization (with a cutoff voltage of 1.00 V). The end-of-charge voltages remain about 30 mV higher than for unsterilized cells. The capacities are stable. The best data are for a pair of cells constructed with the FT 2140 polypropylene separator, 34 wt % KOH, and 80% pore fill. In these cells, the efficiencies based on the nickel plate are 84 and 91%. Cycling is continuing.

The 18-plate cells, which were built with FT 2140, 30 wt% KOH, and 80% pore fill, have been cycled 130 times after sterilization. The delivered capacities remain high and uniform, and the average efficiency for these five cells on the last cycle was 78 $\pm$ 8% of the positive plate formation capacity. Cycling is continuing.

Although the basic designs for the cylindrical and prismatic 25-A-h cells have been established, efforts continue to improve capacity, mid-discharge voltage, and uniformity. The cylindrical cells now provide 14 W-h/lb, an end-of-charge voltage of 1.45 V, and 1.22 V at 50% depth of discharge. In comparison, the prismatic cells yield 13 W-h/lb, an end-of-charge voltage of 1.48 V, and 1.24 V at 50% depth of discharge. A limited production run of the cylindrical cell will be made, and characterization testing will be done.

The seal remains as a major problem area. The two leading candidates are ceramic and crimped-polymeric seals. A great deal of work has been expended in attempts to design a crimped-polymer seal that performs satisfactorily after sterilization, since it has distinct advantages for the impact-resistant cell. Modifications of the structure have been made, and different polymers will be used in an effort to prevent the leaks that have occurred after sterilization.

Initial investigations for the impact-resistant cells have included testing of static constant deformation, static constant load, and dynamic tensile loading. In addition to being used for the design of cells, these data will be used to support a JPL generalized computer program devoted to the computation of nodal deformations and stresses in the various elements of a structure subjected to high-impact loading.

2. JPL Contract With ESB Inc.

This program to develop sterilizable silver–zinc cells is also divided into the two main sections given for the nickel–cadmium cell program. The first section contains developments of an 80-A-h primary cell capable of four 100% cycles, a 10-A-h primary cell capable of four 100% cycles, a 25-A-h secondary cell capable of ninety 60% cycles, and a 48-A-h secondary cell capable of four-hundred 60% cycles. The other section comprises developments of a 5-A-h primary cell capable of four 100% cycles and a 25-A-h secondary cell capable of ninety 60% cycles after sustaining a high-impact shock of 4000 *g*. The status of these developments is as follows:

(1) The 10-A-h primary cells, developed from the 5-A-h, high-impact resistant cell used successfully in the Capsule System Advanced Development Project, have been cycled at 25, 50, 75, and 150% of charge capacity based on a nominal 10-A-h rating. For these cycles, the plateau voltages ranged from 1.41 to 1.54 V, the capacities to a 1.25-V cutoff were 9 to 11 A-h, and the average energy density was 43 W-h/lb. The gassing of H_2 during the formation charge was shown to be due to the use of heat-shrinkable RNF-100 polyolefin tubing; this problem has been eradicated by replacing the polyolefin tubing with heat-shrinkable teflon tubing.

(2) The rechargeable 70 A-h primary cell has been cycled four times. The mean data are: 89 A-h at the C/4 discharge rate, 133 W-h at an average plateau voltage of 1.49 V, and an energy density of 52 W-h/lb. Similar cells have shown a negative plate erosion of about 44%, sticking separator layers, and friable negative plate grids after a 10-mo wet stand and 24 cycles. The cell has been redesigned to include sintered teflonated negative plates to stem erosion, decreased amount of amalgamation to

prevent the loss of grid strength, and framed grids on the positive plate for increased strength.

(3) A factorial experiment is being used to investigate the variables and negative plate density, number of layers of separator, and positive or negative wrapping in efforts to optimize the design for the 25-A-h secondary cell.

(4) In the 400-cycle cell development, some cells have failed as a result of silver and zinc penetration of the six layers of separator after eighty-four 50% cycles and a 6-mo wet stand. These were 24-A-h cells, and the rate of loss of capacity was 0.33 A-h/cycle. A factorial experiment; with the variables of percent of teflon in the negative plate, allowance for thickness per layer of separator, KOH concentration, and ZnO/Ag ratio; is being used to optimize the design.

(5) The present design for the impact-resistant cells utilizes a zircomium sheet in the silver plate and a chemically-etched silver sheet in the zinc plate. The penalties of this design are a reduction in the charge efficiency of about 18% in the silver plate and a reduction of about 7% in the voltage of the zinc plate at the C/2 discharge rate. This effort is being coordinated with the JPL generalized computer program, which is devoted to the computation of nodal deformations and stresses in the various elements of a structure subjected to a high-impact loading.

3. JPL Contract With the Southwest Research Institute

This program is for the development of separator materials made by grafting polyethylene with acrylic acid and crosslinking with divinylbenzene under the influence of Co^{60} radiation. Several parameters involved in this process are being studied. In addition, pilot quantities of the separator material are being made for use in the current research and development work in heat-sterilizable batteries.

The following results have been obtained:

(1) Carbon tetrachloride, acetone, carbon disulfide, and sulfur are good chain terminators for the acrylic-acid grafting of polyethylene.

(2) Aqueous solutions of acrylic acid can be used for grafting polyethylene film if a gelling inhibitor, such as potassium ferricyanide, is used.

(3) An increase in the temperature of the neutralization and washing solutions, used to treat the freshly grafted film, results in a decrease of electrical resistance in the film.

(4) The poly(acrylic acid) content of the grafted film is proportional to the concentration of acrylic acid in the grafting solution. This relation applies to solutions containing from 10 to 25% acrylic acid.

(5) A low carbon-tetrachloride concentration in the grafting solution favors the preparation of materials with low electrical resistance and high silver-oxide diffusion rates, while a higher concentration of the chain terminator results in materials of higher resistance but less rapid silver-oxide diffusion.

4. JPL Contract With Eagle–Picher Industries

This a research and development contract for a remotely activated, heat-sterilizable silver oxide–zinc battery. In the completed first phase, individual components were tested and selected. The second and final phase, which is for the design and fabrication of an engineering model battery, is in progress. The model will be used to determine electrical performance after sterilization.

Various cell designs have been tested. A final design has been selected in which: (1) each cell contains five negative plates and four positive plates; (2) the separation is a sandwich of hemp and asbestos; (3) the cell cases, made from polysulfone, will form an integral part of the battery; (4) the polysulfone is solvent-welded followed by a thermal–vacuum cure of the joint; (5) a copper manifold contains a 31% KOH electrolyte solution; and (6) the heat sterilizable propellant is supplied by JPL. Eagle–Picher Industries has tested various igniters for the propellant and has qualified one. Following a final decision concerning battery design, prototype batteries will be built and tested.

C. Calorimetric Studies of the *Surveyor* Main Battery, *W. L. Long*

1. Introduction

The completion of calorimetric studies of the main *Surveyor* battery represents one immediate objective of a continuing JPL program that seeks to define the performance limits of existing flight batteries. The definition of these performance limits will permit a critical evaluation of the potential usefulness of the existing batteries in the light of requirements imposed by future missions contemplated by JPL/NASA.

In addition, information obtained during the performance of this task that contributes to a more fundamental understanding of both flight battery operation and failure modes is factored into the design of batteries for use on future missions.

This article describes work performed by the Hughes Aircraft Company under JPL contract during the period March 18, 1968 to March 18, 1969 (Ref. 1). The preliminary results of this work were discussed in SPS 37-53, Vol. III, pp. 51–52.

2. Experimental Approach

Figure 10 illustrates the construction of the isothermal calorimeter. It was operated at the boiling point of a

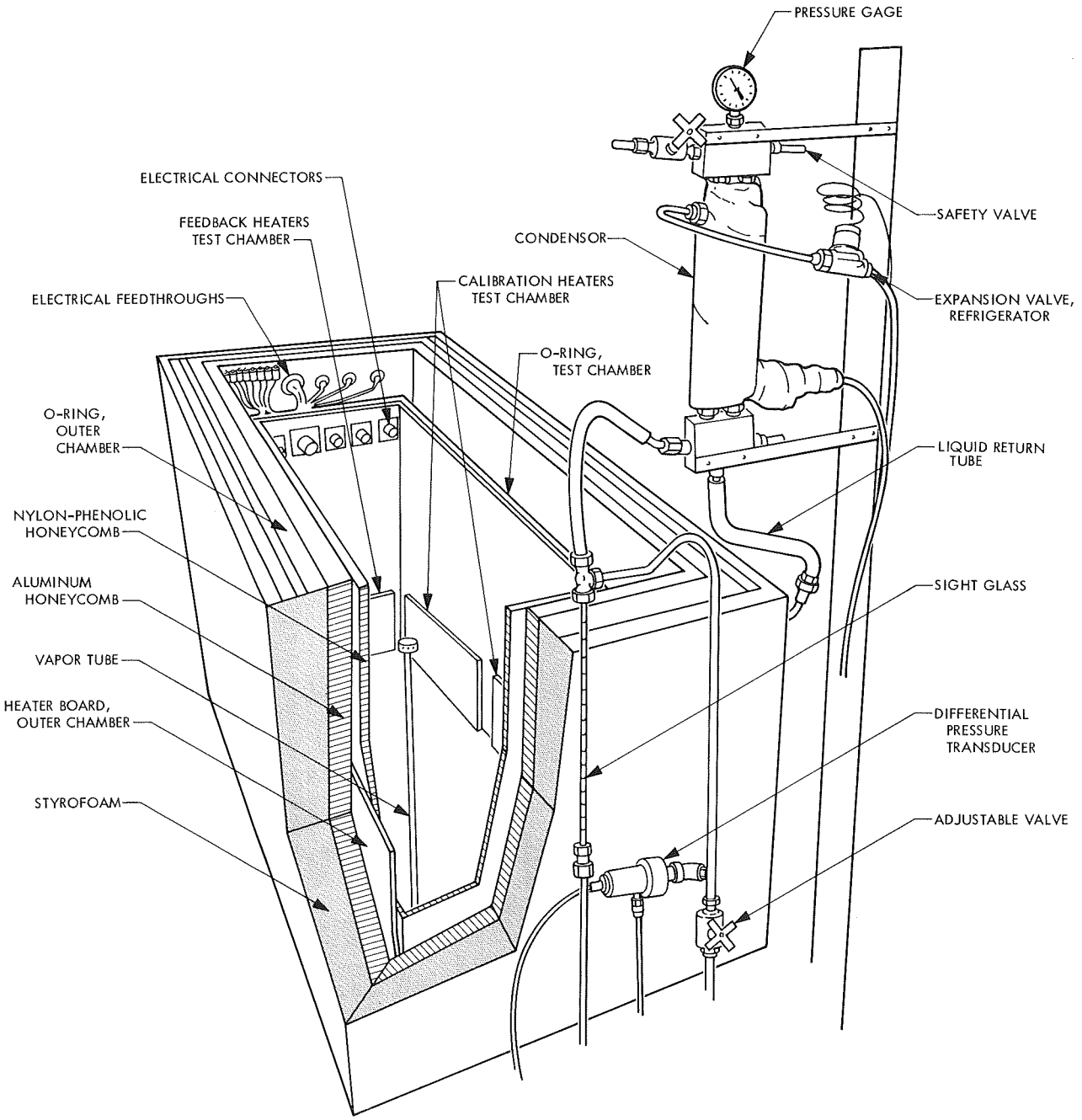


Fig. 10. Isothermal calorimeter

selected freon liquid. A steady state was maintained by means of automatically controlled electric heaters, which kept the boiling rate constant. Figure 11 is a block diagram of the control circuitry. A strip-chart recorder, with suppressed zero, recorded the decrease in electric power supplied to the heaters that balanced the battery heat output.

3. Discussion and Results

As shown in Table 4, heat generation measurements were made on nine batteries. Typical data curves at three different temperatures are shown in Figs. 12, 13, and 14.

Table 4. Battery calorimetric program

Experiment	Charge Current, A	Discharge Current, A	Temperature, °F	Freon
1	0.5–5	0.5–18	75	11
2	—	0.5	75	11
3	—	0.5	75	11
4	2–5	2–18	75	11
5	2–5	2–18	118	TF
6	—	0.5	118	TF
7	2–5	2–18	118	TF
8	2–5	2–18	39	114
9	—	0.5	39	114

Figure 15 shows a battery failure with a large amount of heat generated when discharged at 0.5 A at 39°F. Similar failures occurred with the batteries discharged at 0.5 A at 75°F. The battery discharged at 118°F, 0.5 A load, was discharged completely without failure or excessive heat generation.

It is apparent that the actual heat generated by the *Surveyor* battery during the various charge–discharge profiles follows the calculated heat generation reasonably

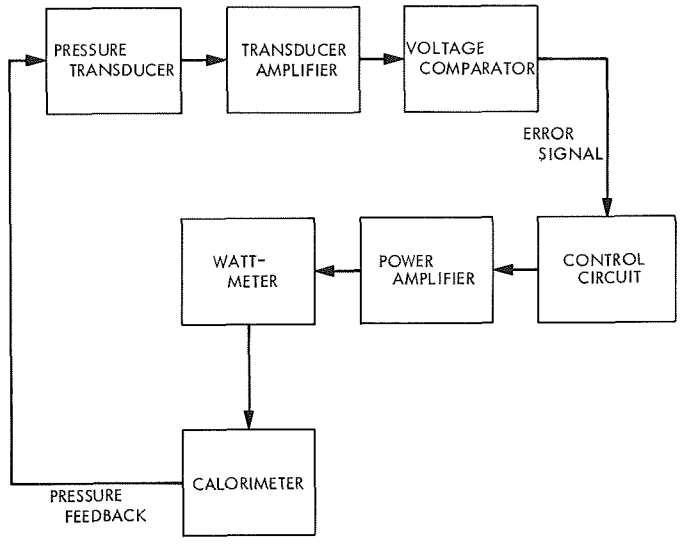


Fig. 11. Calorimeter control block diagram

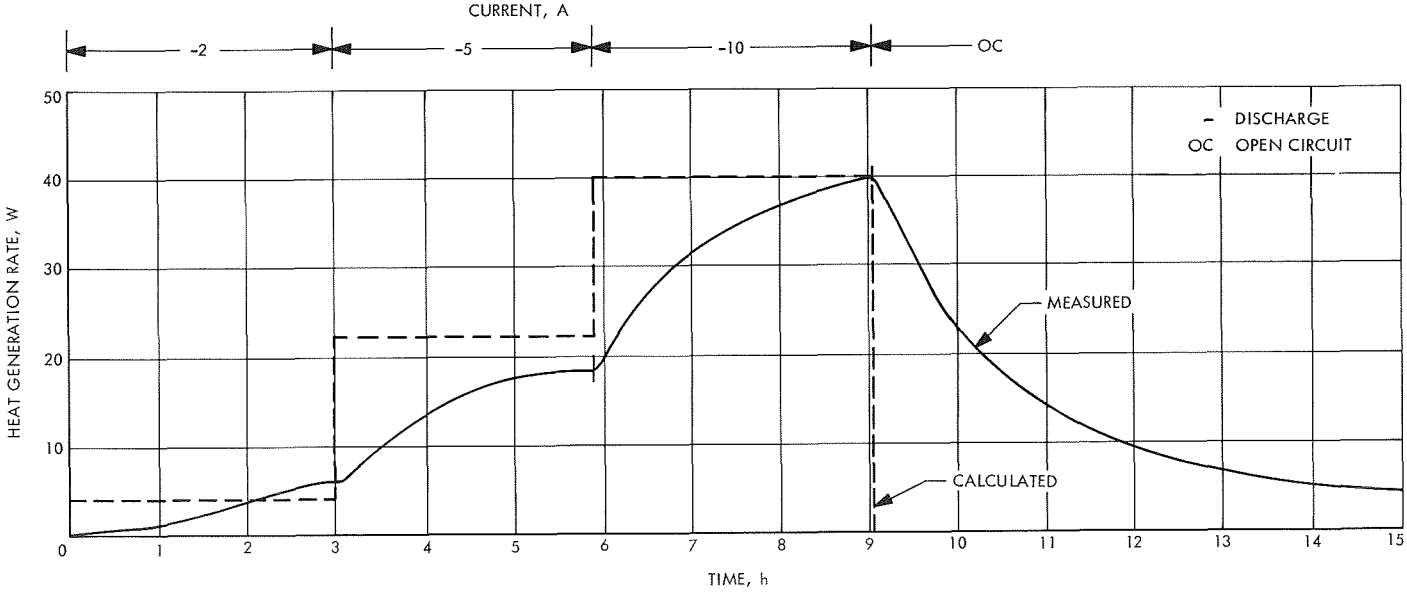


Fig. 12. Heat generation by *Surveyor* battery at 75°F

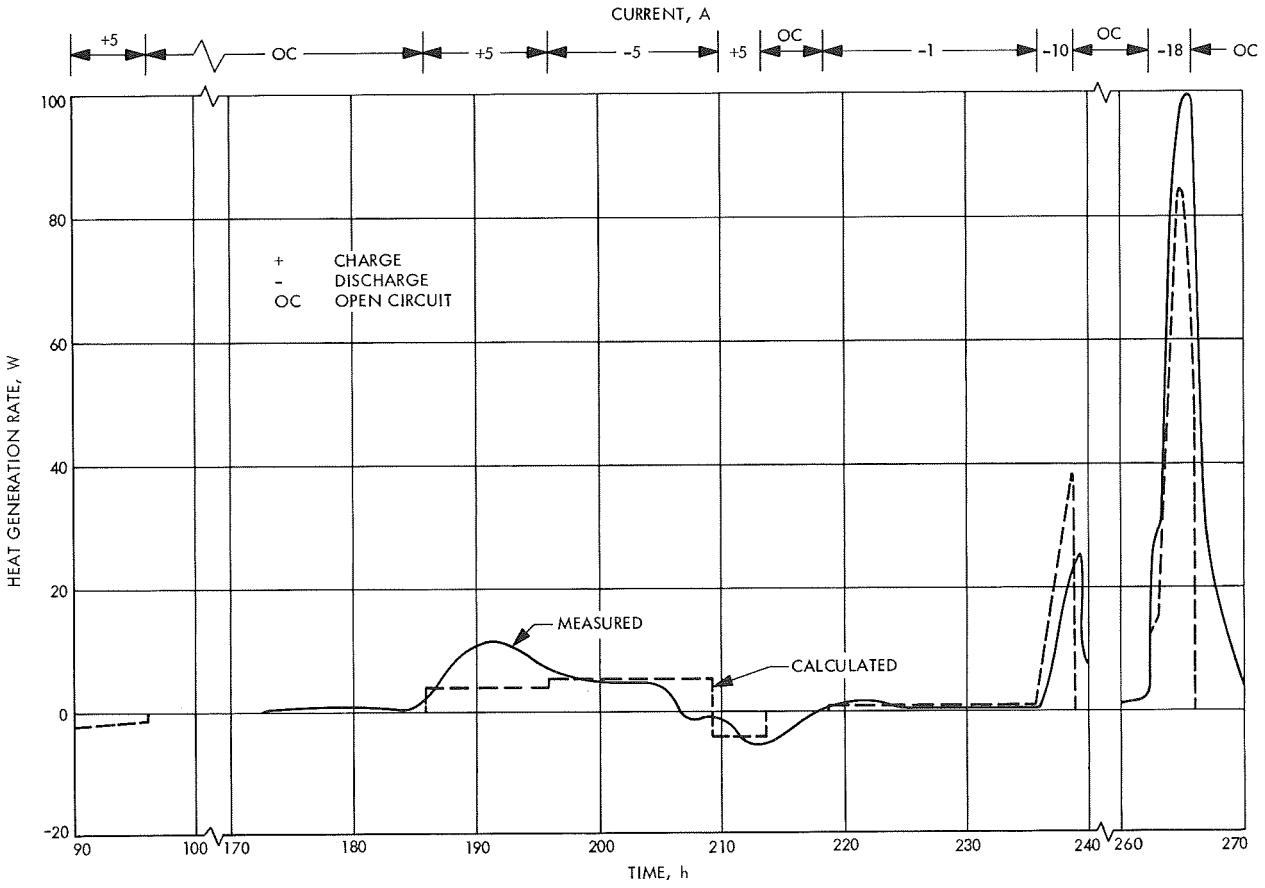


Fig. 13. Heat generation by *Surveyor* battery at 118°F

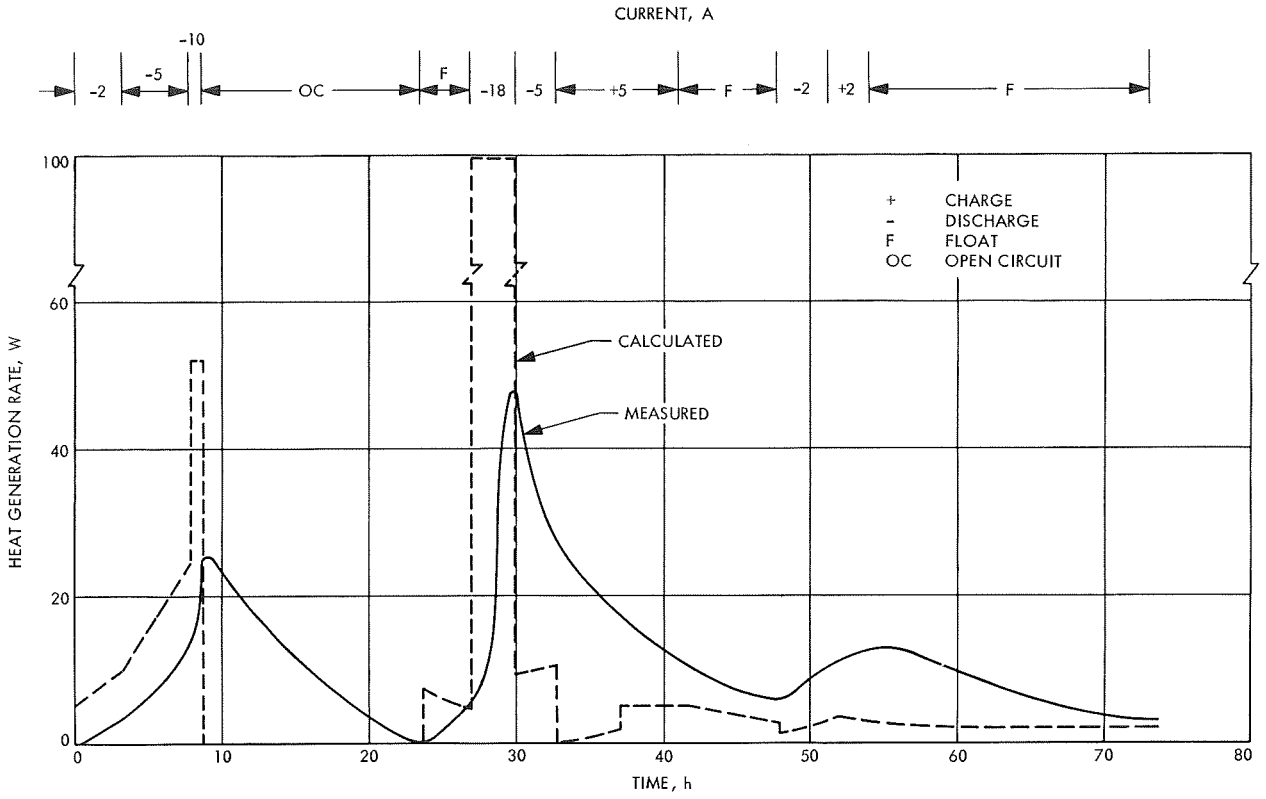


Fig. 14. Heat generation by *Surveyor* battery at 39°F

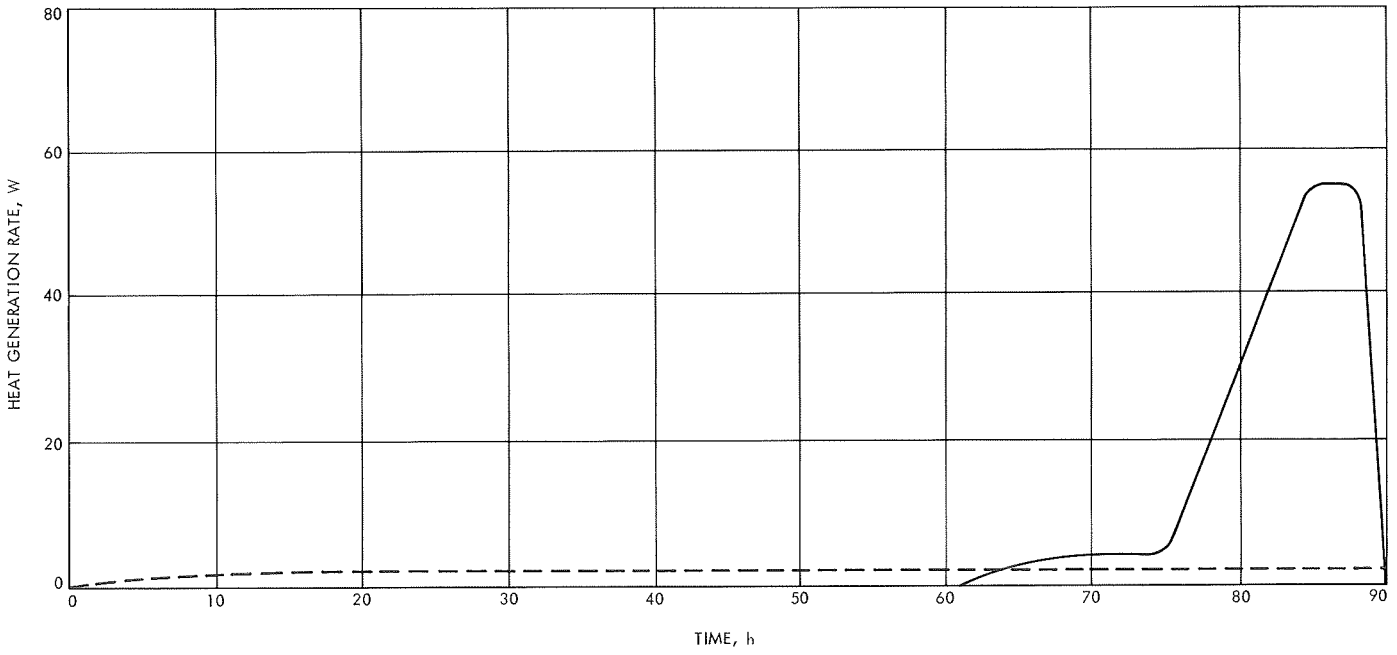


Fig. 15. Heat generation by *Surveyor* battery under 0.5 A discharge at 39°F

well. Consequently, for purposes of thermal control design considerations, the calculated values should be acceptable. In those instances where the measured values are significantly greater, it appears that poor heat transfer characteristics may be playing an important role. The development of spacecraft batteries with better heat transfer properties would represent a significant step forward in battery design.

The effectiveness of the isothermal calorimeter with automatic read-out for direct measurement of the heat generated by spacecraft batteries during various test cycles has been demonstrated.

At the present time, the failures of the batteries discharged at 0.5 A at 39°F and 75°F are not completely understood. Performance of batteries undergoing the same test at 118°F was normal.

Reference

1. Hetherington, W. M., and Weinberger, D., *Surveyor Battery Calorimetric and Cycle Life Tests*, Report SSD 90086R, JPL Contract 952226, Hughes Aircraft Company, El Segundo, Calif., Mar. 1969.

D. Gravity Effects on the Performance of Smooth Zinc Electrodes, *G. L. Juvinall*

1. Introduction

One immediate objective of the JPL program to study the effects of variations of gravitational acceleration on the performance of batteries has been achieved by the fabrication of a breadboard low-gravity battery test unit. The design and construction of a flight unit will commence subsequent to further testing and development of the breadboard unit at JPL.The breadboard unit was fabricated by the General Electric Research and Development Center under JPL Contract 952121.

The capabilities of the breadboard unit include measurement of the limiting current of smooth flooded zinc anodes, measurement of the limiting current and capacity of conventional silver–zinc battery cells, and a photographic study of the formation and behavior of bubbles on the surface of silver and zinc electrodes during flight. The photographic results are correlated with electrical measurements of capacity and limiting current. Details of the cells used for the bubble studies have been presented in SPS 37-52, Vol. III, pp. 38–44.

The subject of this discussion is the measurement of the limiting-current capabilities of smooth zinc electrodes. Earlier work at JPL has investigated the effects of sustained high-g levels in the range of 1–20 g (SPS 37-23, Vol. IV, pp. 20–26; SPS 37-26, Vol. IV, pp. 40–42; and SPS 37-30, Vol. IV, pp. 28–32).

2. Discussion

In the original work at JPL, the limiting current of a 1-cm^2 area of smooth zinc electrode was determined by loading the cell with successively larger discharge currents with each level maintained constant for a time interval of 1 min. This process was terminated when the voltage between the zinc electrode and a zinc reference electrode dropped to a value of 1 V. The maximum current level reached just prior to polarization was defined as the limiting-current density.

In the present method, a linear current–time ramp loading has been substituted. This type of loading gives very reproducible results, which can easily be correlated with the earlier work. In addition, implementation of this experimental approach is much less complicated.

The test cell used in these studies is very similar to that used in the early JPL work. Figure 16 shows two cells, one of which is assembled and the other disassembled. The cell case and cover plate are fabricated from acrylic plastic. The auxiliary electrode employed in the cell case is a silver electrode taken from a Yardney HR5-type silver–zinc cell. The 1-cm^2 electrode area is defined on the zinc test electrode by means of an acrylic plastic mask. The zinc-wire reference electrode is brought in through the main body of the cell.

Each cell is provided with a stem electrolyte port through which it is filled from a common manifold. Electrical isolation of individual cells is accomplished by means of a nylon capillary tube through which the electrolyte must enter. This tube may be seen in Fig. 16 coiled around the cell body. The dimensions of this tube are such as to give a minimum resistance of 10^4 Ω when it is filled with electrolyte. The electrolyte is 40% potassium hydroxide saturated with zinc oxide.

The cells, which are mounted in a bank of six in the breadboard unit, may be oriented in either a vertical or horizontal position with the test electrodes parallel or perpendicular to the gravity vector, respectively. Although a valid simulation of 0 g conditions is not feasible in the

laboratory, the dependence of the limiting current capability of the zinc electrodes on the orientation and magnitude of gravitational acceleration may be demonstrated.

3. Results

A sample chart record for one cell tested in a vertical configuration is shown in Fig. 17. Here, on the graph of cell current versus time (Fig. 17c), the end of the ramp clearly shows that the limiting current has been reached (Fig. 17b), and that polarization of the zinc electrode has occurred (Fig. 17a).

A summary of the data obtained on cells tested in the vertical configurations is shown in Table 5. These data illustrate the very good reproducibility of the method. A comparison of these results with the data from cells tested in the horizontal configuration shown in Table 6 reveals the sharp difference in performance of horizontal and vertical electrodes. In the case of the horizontal electrode, a relatively stable dense layer of electrolyte rich in zincate ion is built up at the bottom of the cell during the discharge of the zinc plate. If the electrode is in a vertical position, convection will tend to sweep the reaction products away from the reacting surface, thus contributing to the mass transfer process and increasing the limiting current capability of the electrode.

At 0 g, the cell orientation would not be expected to influence the performance of the electrodes; the mass transfer process would then be diffusion controlled. The

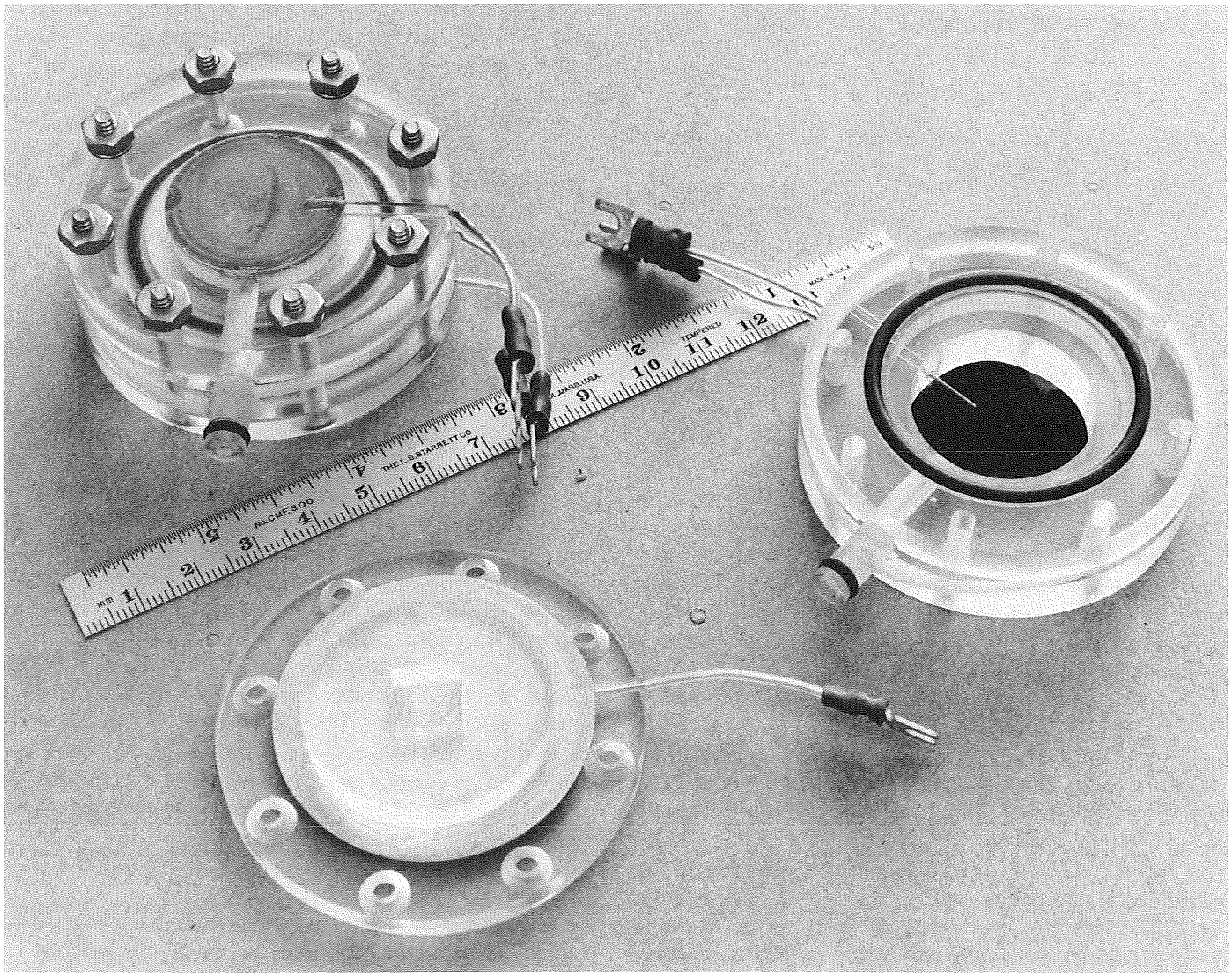


Fig. 16. Assembled and disassembled test cells

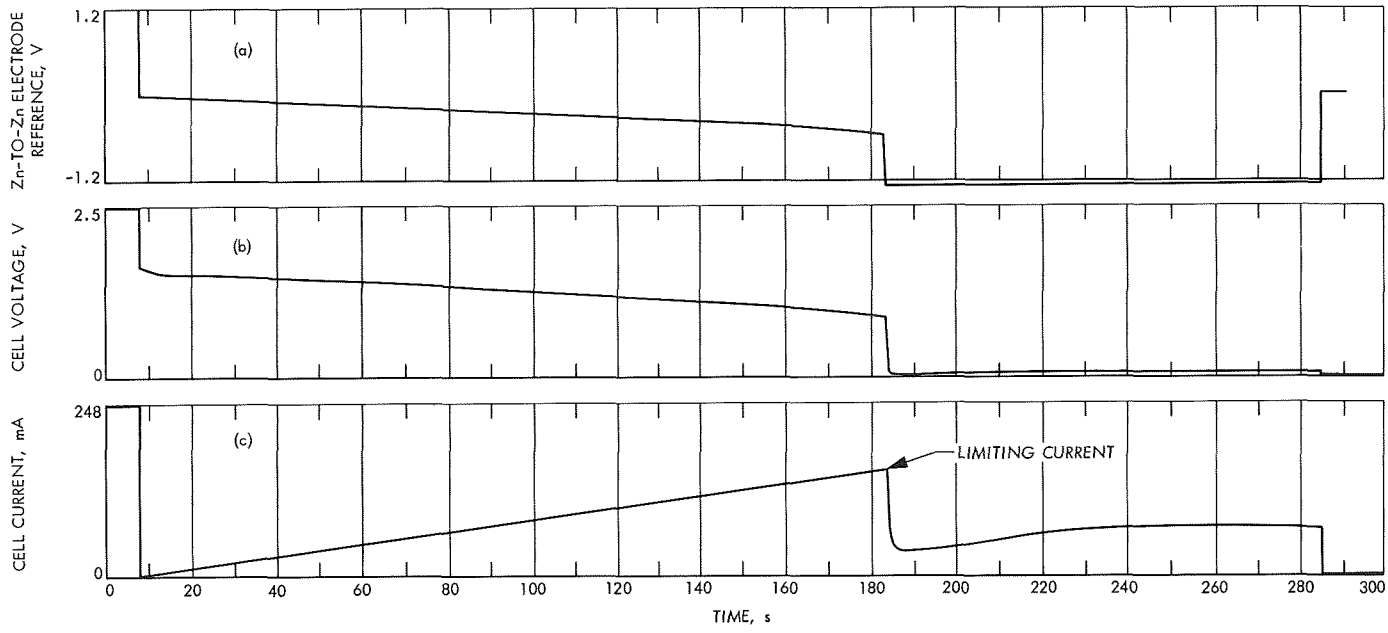


Fig. 17. Vertical configuration test chart record

performance of these measurements in space by means of a flight experiment will lead to a better understanding of the performance limits of flight batteries.

Table 5. Vertical electrode position

Test cell	Limiting current density, mA/cm²	Time to polarization, min	Discharge capacity, mA–min
1	141.4	2.68	190
2	139.0	2.60	179
3	149.0	2.80	207
4	156.2	2.93	227
5	141.3	2.68	190
6	151.3	2.85	215
Mean	146.37	2.757	201.3

Current ramp slope: 52.8 mA/min
Temperature: 24°C

Table 6. Horizontal electrode position

Test cell	Limiting current, density, mA/cm²	Time to polarization, min	Discharge capacity, mA–min
1	65.8	9.3	295
2	65.8	9.3	295
3	69.5	10.1	342
4	69.5	10.2	354
5	65.8	9.5	306
6	65.8	9.7	320
Mean	67.03	9.68	318.7

Current ramp slope: 6.83 mA/min
Temperature: 24.5°C

E. Charge Acceptance of Silver Electrodes Under Conditions of Controlled Potential, *G. L. Juvinall*

1. Introduction

The acquisition of basic knowledge relevant to the fundamental processes involved in the operation of spacecraft batteries is one major objective of the battery advanced development effort at JPL. In support of this effort, studies of the electrode surface reactions involved in the oxidation and reduction of the alkaline silver electrode are being performed at Brigham Young University under JPL Contract 952268. Dr. Eliot Butler is the principal investigator. Other work included in this task has been discussed in SPS 37-39, Vol. IV, pp. 19–22; SPS 37-43, Vol. IV, pp. 66–69; SPS 37-53, Vol. III, pp. 47–49; and SPS 37-56, Vol. III, pp. 119–120.

2. Discussion

The fundamental importance of the potential in the electrocrystallization of oxides on metal electrodes suggests that potentiostatic oxidations should be of special value in studies of the mechanism of the oxidation of silver electrodes in alkaline solution. In addition, data on the total charge acceptance as a function of applied potential would have practical application in the areas of battery system and charger design.

In Ref. 1, P. A. Malachesky and R. Jasinski have presented data on the total charge acceptance as a function

of the applied potential for argentous and argentic oxide. The potential range spanned the region between the values required for formation of argentous oxide and the evolution of oxygen, respectively.

A plot of total charge acceptance versus applied potential over the same potential range used by Malachesky and Jasinsky was obtained in the Brigham Young University Laboratories. Although it has the same general shape as the plot of Malachesky and Jasinski in the argentous and argentic oxide regions, a third reaction at potentials just below the required value for the evolution of oxygen was found to yield a large charge acceptance peak, which was not reported by Malachesky and Jasinski.

The electronic circuitry used for the potentiostatic oxidations is shown in Fig. 18. Definitions of the letter designations used in Fig. 18 are given in Table 7. The electrolyte was a 0.1 N potassium-hydroxide solution that had been saturated with argentous oxide to prevent dissolution of oxide from the electrode. The reference electrode was a mercury–mercuric-oxide electrode. All the oxidations were carried out at $20 \pm 0.01°C$ and both abraded-foil and wire electrodes were used. The current and integral of the current were recorded for all oxidations. The quantity of silver that had been oxidized was determined using standard atomic absorption techniques subsequent to stripping the oxide film from the electrode by means of a 0.5 N solution of ammonium hydroxide.

3. Results

Figure 19 gives the results of a comparison of the total charge acceptance as determined by the current–time integral (curve A) and the amount of silver oxidized

Table 7. Definitions of symbols used in Fig. 18

Symbol	Definition
A	Potentiostat—Philbrick P25AU operational amplifier
B	Booster—Philbrick P66A power amplifier (100-mA mode)
C	Integrator—Philbrick P25AU operational amplifier
D	Current recorder
E	Constant potential power supply
F	Integrator output recorder
G	100-Ω standard resistor
H	Hg–HgO reference electrode
I	Ag working electrode
J	Pt counter electrode

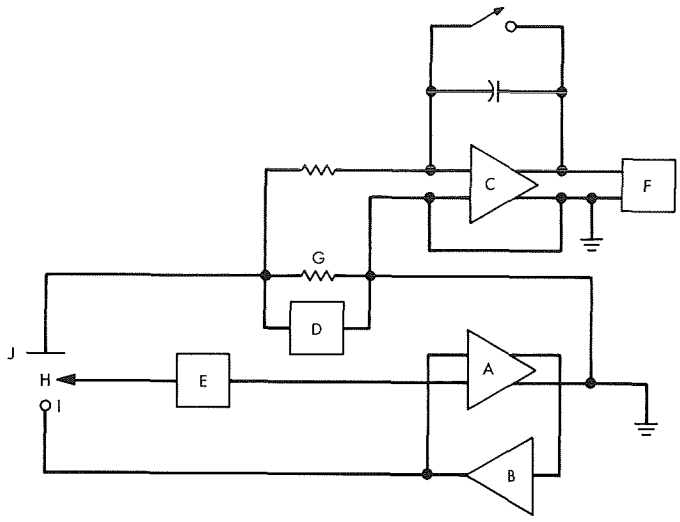


Fig. 18. Circuit of potentiostat with 100 mA booster amplifier

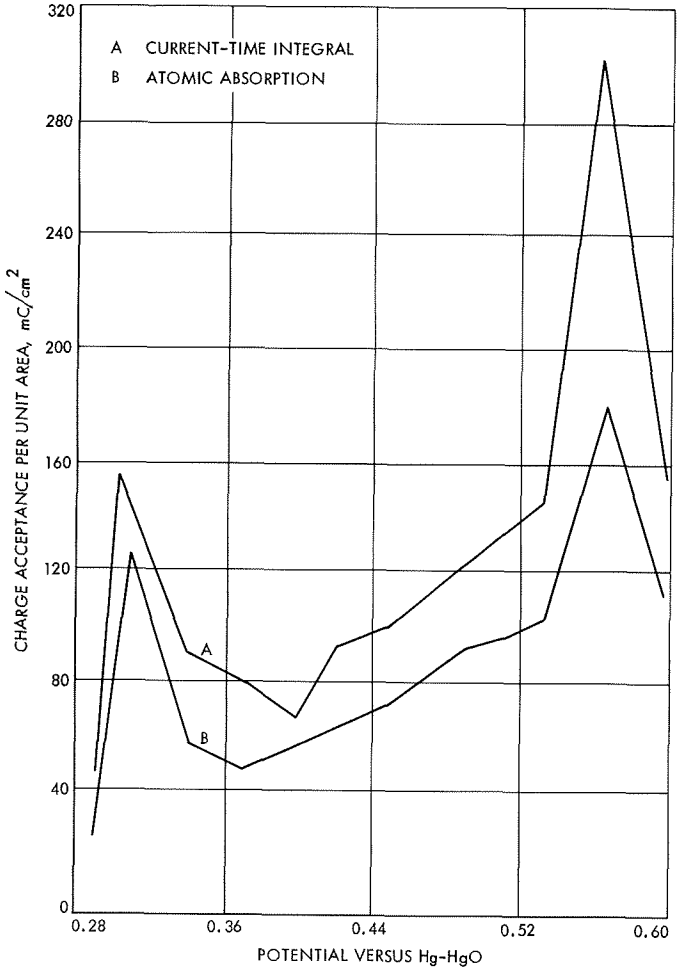


Fig. 19. Charge acceptance versus applied potential

(curve B) as functions of the applied potential. The significant point to be noted in this comparison is that both curves show the same maximum–minimum relationships. The fact that the current–time integral is consistently higher indicates either a current efficiency of less than 100% or loss of the silver oxide from the electrode. However, the data provide unequivocal evidence of the wide range of charge acceptance capabilities that may be exhibited by the silver electrode under potentiostatic conditions.

Future work will include attempts to determine if the initial thin layer of silver oxide determines the subsequent behavior of the electrode. It has been postulated (Ref. 2) that the extent of the reaction on the electrode surface is limited by the migration of ions through this thin layer, which represents only about 20 mC/cm^2 of the total charge acceptance. Clarification of this point is necessary in order to better understand the chemical reactions and processes occurring at the electrode–electrolyte interface in alkaline batteries.

References

1. Malachesky, P. A., and Jasinsky, R., Tyco Lab Inc., Waltham, Mass.; ABS 174, Electrochemical Society Inc., Spring Meeting, Boston, Mass., May 5–9, 1968.
2. Fleischmann, M., Lax, P. J., and Thirsk, H. R., *Trans. Faraday Soc.*, Vol. 64, No. 3128, 1968.

F. Simulation of a Pu^{238} RTG Fuel Capsule, *M. Reier*

Detectors contained in a typical science package for outer-planet mission spacecraft may be sensitive to the radiation environment produced by a radioisotope thermoelectric generator (RTG). It is therefore important to measure the response of these detectors to such an environment. Although it is not feasible to keep an RTG source in the radiation laboratory for extended periods of time, the gamma-ray radiation spectrum, which is of prime importance when considering interference, can be simulated by the appropriate grouping of small quantities of a number of different isotopes that are primarily gamma emitters. Therefore, such radioactive sources have been ordered that together will simulate the gamma-ray spectrum and intensity from the fuel and impurities in a *Systems for Nuclear Auxiliary Power* (SNAP) 27 heat source producing 1575 W of thermal power. It is anticipated that the radiation from the simulated source will be compared with an actual SNAP 27 heat source for purposes of verification of the calculations. Minor adjustments will be made in order to achieve the desired duplication. Table 8 shows the composition of the SNAP source.

The simulated source includes the gamma radiation from the decay of the Pu isotopes and their daughter products, as well as gamma rays resulting from the (α, n) reaction in O^{17} and O^{18} since the fuel is processed as PuO_2.

Table 9 shows the quantities, in millicuries, of radioisotopes required to duplicate the SNAP 27 spectrum as a function of the number of years after processing of the fuel having an original composition as shown in Table 8.

The Th^{228}, which builds up from the decay of Pu^{236}, ultimately decays to Tl^{208}. This emits numerous gamma rays, the most important of which (from the standpoint of shielding difficulty) has an energy of 2.62 MeV.

The source that has been purchased for use in the RTG radiation test laboratory corresponds to the 5-yr fuel sample shown in Table 9. A source equivalent to 18-yr fuel will also be purchased in the near future. It is relatively easy, however, to mock up other years by proper geometric configurations or judicious time exposures with the individual sources. Freshly processed fuels can be simulated by removal of the Th^{228} source.

Plates of U^{238} will be used to simulate the self-absorption of the fuel and the shielding by the cladding. Transport calculations already performed have shown that this can be done with a high degree of accuracy.

Table 8. Relative abundances of Pu isotopes in the SNAP 27 source

Isotope	Abundance, %
Pu^{238}	81
Pu^{239}	15
Pu^{240}	2.9
Pu^{241}	0.8
Pu^{242}	0.1
Pu^{236}	1.2×10^{-4}

Table 9. Composition of simulated source

Year	Th^{228}, mCi	Ba^{133}, mCi	Cs^{137}, mCi	Zn^{65}, mCi
0	0	0	13.3	2.20
5	38.1	1.14	13.0	2.12
18	84.8	1.46	12.0	1.91

G. Effects of Pu^{236} on Shielding Requirements for RTG Spacecraft Experiments, M. A. Dore

1. Introduction

Of major concern in the design of spacecraft for exploration of the outer planets is the interference of the gamma radiation field from the radioisotope thermoelectric generators (RTGs), used as a power source, with sensitive scientific instruments, particularly charged-particle detectors. The extent of such interference is a strong function of the number and energy of the gamma rays reaching these sensitive detectors. For Pu^{238}, the most probable isotope fuel for such missions, the number and energy distribution of the emitted gammas are highly dependent upon both the lapsed time since chemical separation of the fuel and the relative concentration of the impurity isotope Pu^{236} immediately after such separation. Pu^{236} has a 34% yield of Th^{208}, which in turn decays to Pb^{208}, emitting a 2.615 MeV gamma 100% of the time and giving rise to a very hard, time-dependent spectrum.

Pu^{236} is formed from $(n, 2n)$ and (γ, n) reactions during Np^{237} irradiation. S. V. Topp (Ref. 1) has indicated that the amount of this impurity isotope depends on the type of reactor used to irradiate the Np^{237}. Light-water reactors produce Pu^{238} containing 20 to 40 times as much Pu^{236} as that produced in the AEC D_2O or graphite production reactors. The liquid metal fast breeders produce 600 to 900 times as much Pu^{236} as the production reactors. To date, application of Pu^{238} to space power has been from fuel produced in the production ion reactors, which contains about 1.2 parts/10^6 of Pu^{236}. By letting the fuel decay for a few years ($t/2$ of 2.85 yr for Pu^{236}) and then re-processing it, it is possible to achieve 0.1 parts/10^6 or so (called "biomedical grade").

2. Basic Assumptions

a. RTG fuel capsule. The design chosen here for the capsule contains 1575 W(th) of fuel, assuming the available space is 61.6% filled by the PuO_2 microspheres. The fuel is contained in an annular cylinder of tungsten, which comprises about 50% of the capsule volume, the remainder being primarily void volume to allow for the build-up of helium from the decays. The entire capsule is then surrounded by a tantalum casing, forming a right-circular cylinder about 40 cm long and about 9 cm in diameter. Assuming a 5% efficiency for the generator, this capsule would give rise to about 80 W(e), which is typical for the per-unit wattages anticipated for outer planet spacecraft RTGs.

b. Source and spectrum. Spectra were determined on the basis of 20 energy groups, ranging in energy from 1 KeV to 7 MeV, as shown in SPS 37-56, Vol. III, pp. 128–132. Source terms are due to P. J. Gingo (SPS 37-56, Vol. III)[1]. The communication referenced in footnote 1, which contains modifications and corrections to the previous compilation, was the one actually used. Nominal fuel composition resulting from graphite production reactors was assumed for the initial computer runs and divided into seven distinct sources as follows:

(1) Pu^{238} + daughters, fission products, and minor impurities.

(2) O^{18} (α, n) Ne^{21}.

(3) Pu^{236} + daughters: 0 yr after separation (1.2 parts/10^6).

(4) Pu^{236} + daughters: 1 yr after separation (1.2 parts/10^6).

(5) Pu^{236} + daughters: 5 yr after separation (1.2 parts/10^6).

(6) Pu^{236} + daughters: 10 yr after separation (1.2 parts/10^6).

(7) Pu^{236} + daughters: 18 yr[2] after separation (1.2 parts/10^6).

These sources are run independently to determine fluxes exterior to the capsule, which can then be adjusted according to the relative source strength of each term and/or the capsule age and added together to yield a final spectrum. Some representative spectra are shown in Fig. 20.

c. Geometry. In an actual RTG, the fuel capsule is surrounded by the generator, composed primarily of either lead–telluride or silicon–germanium alloys, arranged in dozens (or even hundreds) of individual thermoelectric cells, each containing several kinds of heavy metals as well. Such a generator has not been included for this study due to both its extreme geometrical complexity and the significantly different gamma attenuation properties of the alternative thermoelectric materials under consideration for use in actual RTGs. This subsection therefore describes a worst-case situation, assuming no attenuation whatsoever from the generator.

[1]Also, Gingo, P. J., and Dore, M. A. *Gamma Radiation Characteristics of Plutonium Dioxide Fuel*, (private communication).

[2]Worst case.

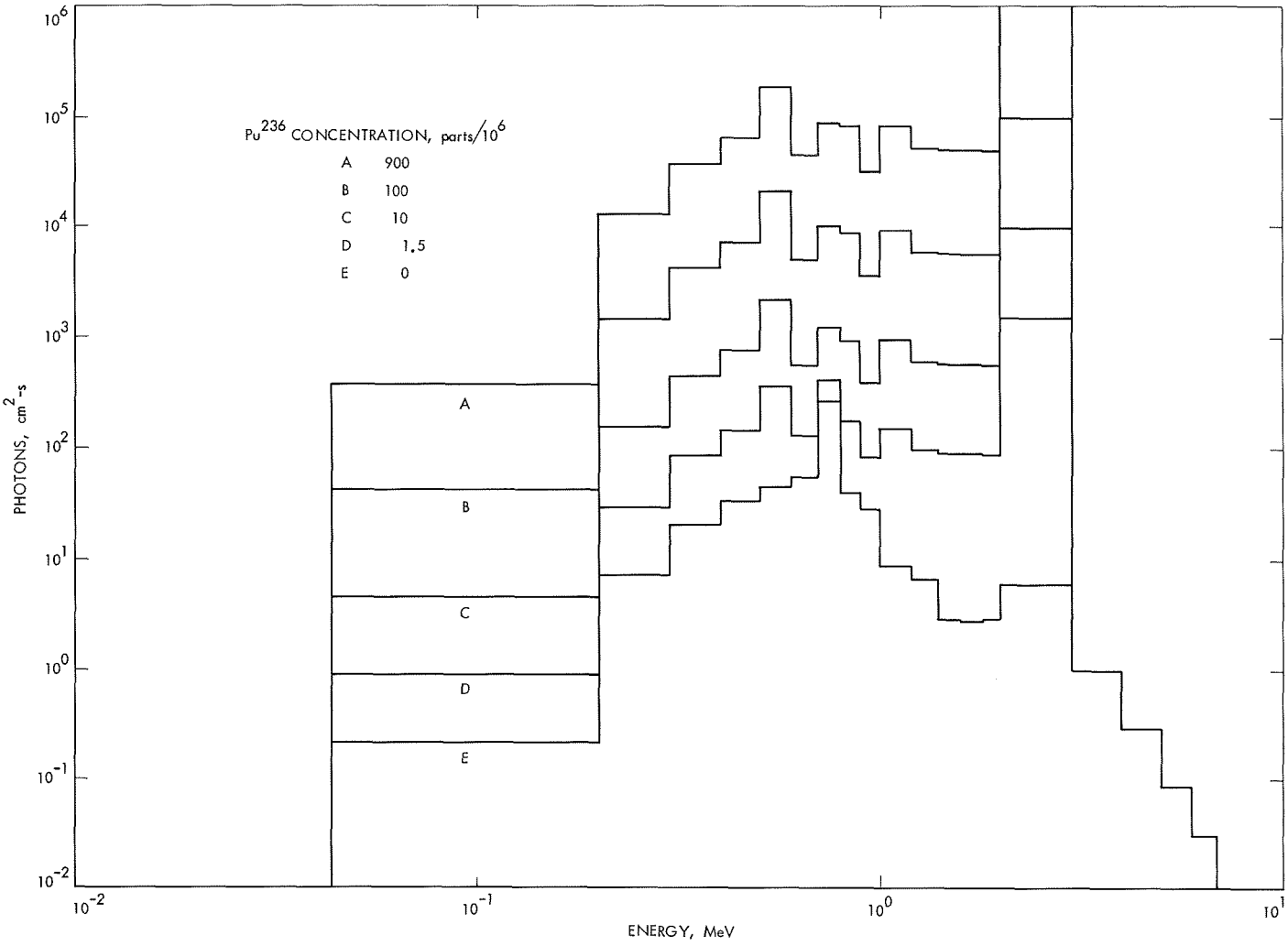


Fig. 20. Total spectrum at 10 yr

The experiment package containing the sensitive instruments is assumed to be located 10 ft from the capsule, which represents a reasonable minimum separation distance for this type of spacecraft. This package is oriented so as to lie in the capsule midplane (again, as a worst-case), the effects of self-shielding within the capsule being thereby minimized.

d. Detector response, shield attenuation, and computer codes. Assumptions and usage here are identical to the material presented previously in SPS 37-57, Vol. III, pp. 68–73. As detailed therein, two codes (QAD-B and ANISN) are utilized to generate flux maps for 20 group spectra, five fuel ages, and any desired points in space. This information can now be utilized by a special-purpose computer code to produce the results presented here.

e. Detectors. A typical complement of science experiments has been selected for this investigation. These include the following:

(1) A low-energy proton and electron differential energy analyzer (LEPEDEA) experiment containing one channeltron electron multiplier (CEM-4010) and one Geiger–Muller tube (EON-6213).

(2) A UV photometer containing one channeltron and one Geiger–Muller tube.

(3) A trapped-radiation detector containing three Geiger–Muller tubes and one solid-state detector (12 mm^2 × 31 μm).

(4) A cosmic-ray telescope containing a single solid-state detector (5.7 cm^2 × 200 μm).

This is the same package described in SPS 37-57, Vol. III, and the assumptions as to sensitive area, energy response, etc., are as described therein. (See also Table 10.) Similarly, shielding is treated as in SPS 37-57, Vol. III and need not be repeated here.

3. Calculational Technique

The mathematical methods employed for the calculations are elementary and straightforward. Letting $j(=1,2,3,4)$ index the detectors (i.e., Geiger–Muller tube, Channeltron, etc.), and letting $k(=1,2,3,4)$ index the experiments (i.e., LEPEDEA, UV Photometer, etc.), the spurious count rate in detector j of experiment k is given by:

$$C_{jk}(t) = A_j \int_0^\infty \xi(E,t)\,\eta_j(E)\,S(E)\,dE$$

$$\cong A_j \sum_{i=1}^{20} \xi_i(t)\,\eta_{ij}\,S_i \qquad (1)$$

where

A_j = sensitive area of detector j

$\eta_j(E)$ = relative energy response of detector j

$\xi(E,t)$ = relative shield attenuation at energy E, thickness t

$S(E)\,dE$ = gamma number flux/s-cm² in energy width dE at the surface of the detector

The code finds the sum, where i indexes the 20 energy groups. This sum is determined by the code for all actual combinations of detectors and experiments for shield thicknesses from 0 to 7 in. in ¼-in. increments.

For each actual combination, the code then interpolates to find the thickness T_{jk} such that

$$C_{jk}(T_{jk}) = N_{jk} \qquad (2)$$

where N_{jk} is the maximum allowable spurious count rate for detector j in experiment k.

Shield weight for each individual detector is then simply the product of this thickness, the shield area, and the density of tungsten. These are summed, first by experiment (including the number of each type of detector in the particular experiment), and then for the entire spacecraft science package, yielding the total required shield weight.

Table 10. Summary of results (all values are for 10-yr old fuel)

Experiment			LEPEDEA		UV photometer		Trapped-radiation detector		Cosmic-ray telescope
Detector			Channeltron	Geiger-Muller tube	Channeltron	Geiger-Muller tube	Solid-state 12 mm² × 31 μm	Geiger-Muller tube	Solid-state 5.7cm² × 200μm
Sensitive area, cm²			2.0129	0.9484	2.0129	0.9484	0.7330	0.9484	8.4800
Maximum allowable spurious counts/s			0.1	0.3	0.1	0.3	0.001	0.3	1.0
Flat shield area, cm²			6.0	4.0	10.0	4.0	3.0	2.0 Each	10.0
Shield thickness, in.	Pu^{238}, parts/10^6	0.1	1.758	0.950	1.758	0.950	1.238	0.950	0.384
		1	2.700	1.840	2.700	1.840	1.984	1.840	0.859
		10	3.795	2.899	3.795	2.899	3.068	2.899	1.813
		500	5.896	4.867	5.896	4.867	5.156	4.867	3.729
Flat shield weight, lb		0.1	1.005	0.362	1.674	0.362	0.354	0.543	0.366
		1	1.543	0.701	2.571	0.701	0.567	1.051	0.818
		10	2.169	1.104	3.614	1.104	0.877	1.657	1.727
		500	3.369	1.854	5.615	1.854	1.473	2.781	3.551
Shield weight per experiment, lb		0.1	1.367		2.036		0.897		0.366
		1	2.244		3.272		1.618		0.818
		10	3.273		4.718		2.534		1.727
		500	5.223		7.469		4.254		3.551
Total Shield weight, lb		0.1			4.665				
		1			7.952				
		10			12.252				
		500			20.499				

This entire process is then repeated 100 times for five different spectra (corresponding to the five fuel ages considered) and for 20 different initial concentrations of Pu^{236}. The final results are presented in Table 11 and Fig. 21.

4. Results

The large amount of information generated as described in this article allows a number of observations and conclusions, particularly as regards practical application to highly restricted spacecraft environments. Assuming the maximum tolerable spurious count levels indicated in Table 10, the shield weights given in Table 11 and Fig. 21 result. As shown, the weights vary logarithmatically with Pu^{236} concentration above about 3 parts/10^6 and approach a constant minimum value below this concentration. For example, shield weights of about 8½ lb are required for 1.2 part/10^6 18-yr-old fuel. Reduction of initial Pu^{236} concentration to the 0.1 part/10^6 biomedical grade (which would almost certainly be financially prohibitive) yields a 5-lb required shield weight (a savings of only 40%), which appears far too small to justify use of the higher quality fuel. Liquid–metal breeder reactor fuel, however, with typically 500 parts/10^6 Pu^{236}, causes almost 21 lb of shielding to be used for 10-yr old fuel. This is definitely a significant increase, particularly for geometrical considerations, since the shield thickness here is about 6 in. and would probably not be compensated for by the slight decrease in fuel cost.

Table 11. Total tungsten spot shield weight for science package

Pu^{236} fraction, parts/10^6	Fuel age, yr				
	0	1	5	10	18
0.0	3.122	3.122	3.122	3.122	3.122
0.1	3.122	3.230	4.170	4.665	4.814
0.5	3.122	3.665	5.929	6.804	7.022
1.0	3.122	4.030	6.995	7.952	8.179
1.5	3.122	4.385	7.647	8.644	8.907
2.0	3.122	4.627	8.150	9.177	9.442
3.0	3.123	5.089	8.874	9.924	10.189
4.0	3.123	5.421	9.409	10.484	10.733
5.0	3.123	5.707	9.806	10.890	11.160
6.0	3.123	5.959	10.151	11.256	11.530
7.0	3.124	6.190	10.460	11.560	11.815
8.0	3.124	6.389	10.698	11.806	12.062
9.0	3.124	6.576	10.916	12.022	12.322
10.0	3.125	6.738	11.120	12.252	12.534
15.0	3.126	7.380	11.910	13.033	13.339
20.0	3.127	7.884	12.500	13.639	13.900
50.0	3.136	9.522	14.335	15.522	15.791
100.0	3.150	10.819	15.758	16.946	17.262
500.0	3.250	13.987	19.215	20.499	20.798
900.0	3.335	15.200	20.531	21.819	22.137

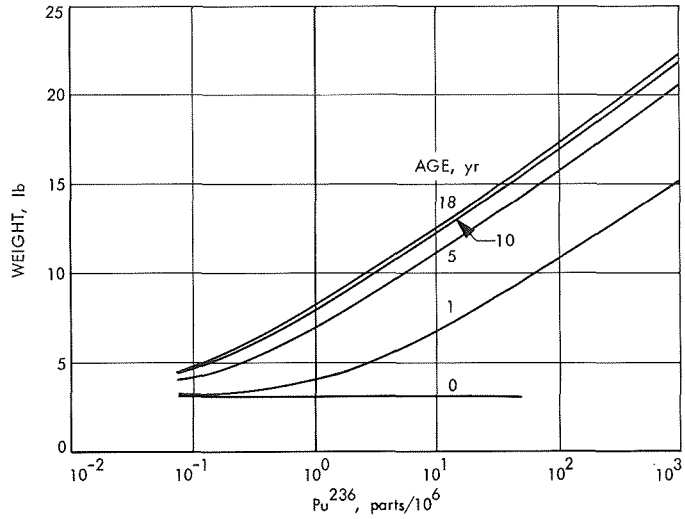


Fig. 21. Total tungsten spot shield weight for science package

Table 10 presents a summary of the most significant results in typical and extreme cases. Many of these values have changed a number of times as this study has progressed, and will probably continue to change as better and more exacting refinements are applied to the problem.

Reference

1. Topp, S. V., "Pu-236 Contaminant in Pu-238 Produced in Power Reactors," *Trans. Am. Nuc. Soc.*, Vol. 11, p. 457, 1968.

H. Small Cylindrical Converters, *P. Rouklove*

1. Introduction

The thermionic energy converter development effort, which has been successfully pursued at JPL since 1962, has been redirected from the emphasis on planar electrode converters, developed for the solar-electric thermionics (SET) program, to cylindrical converters. These devices are being developed for use with nuclear energy produced by reactors or by isotope decay. This article presents the work pursued in the development of small cylindrical thermionic converters capable of operating with the heat released by the decay of an isotope.

2. System Studies

The results of numerous experiments on thermionic energy conversion devices, which were performed in this laboratory and elsewhere, indicate that the ignited mode

operation leads to thermionic conversion efficiencies of 10% and specific weights of about 3–4 W/lb with radiator temperatures of about 1000°K.

These high efficiencies and specific weights promote interest in the use of isotopically heated thermionic generators for use in missions in which the distance from the sun, the environment, the weight management, and the spacecraft design renders the use of photovoltaic or radioisotope thermoelectric power supplies difficult.

It was realized that achievement of high reliability for the overall system was required and several studies to that effect were undertaken (Refs. 1–4). At JPL, the problem was pursued further, and a study was performed (SPS 37-49, Vol. III, pp. 115–117) to outline the optimum size of thermionic converters to be utilized in the most efficient way in power systems of 50–500-W electrical output. Results of these studies indicated that a small converter with a specific power density of 4 W/cm² and a total power output of 15–20 W is the most appropriate to cover the widest range of applications. From a preliminary system study, the emitter operating temperature was selected arbitrarily as 1400°C. It is felt that isotope capsules capable of operating at these temperatures will be developed in the not too distant future. To concentrate the low thermal density, generally associated with long-life isotopes, it was proposed to assemble the converters around a "heat pipe" taking advantage of the well-known and demonstrated thermal concentration properties of these devices.

Analysis of the technology of this "system," and previous experience with the SET type converters, indicated two areas requiring further technical improvements (the problem of developing an acceptable heat source was left for future efforts). The two areas are: (1) the thermionic converter and (2) the integration of the converters with the isotope capsule. Of the two, the highest priority was assigned to the thermionic converter.

3. Converter Development

The Electro-Optical Systems branch of the Xerox Corporation was assigned (JPL Contract 952255) to develop a small cylindrical converter fulfilling the following requirements:

(1) Minimum power output of 16 W at an emitter temperature of 1400°C.

(2) Minimum power density of 4 W/cm² at a voltage output of 0.5 V at the leads.

(3) Minimum size and weight compatible with an emitter area of 4 cm².

(4) Maximum reliability.

This effort, of 18-mo duration, is of an iterative nature. Each converter assembled is to be evaluated by both the manufacturer and JPL and desirable improvements are to be implemented in the next converter to be fabricated.

4. Results

The first cylindrical converter, designated SC-1, was assembled using a bulk rhenium emitter and a vapor-deposited rhenium collector surface. The rhenium was deposited on niobium to reduce the thermal stresses in the metal–ceramic seal. The SC-1 converter (Fig. 22a) was successfully assembled and processed. During the operational tests, however, a short circuit between the emitter and collector developed. Analysis of this deficiency indicated the possibility of an "oil canning" effect of the thin emitter support membrane.

A second converter (SC-2) was assembled. In this device, an alumina ring separator was placed on the collec-

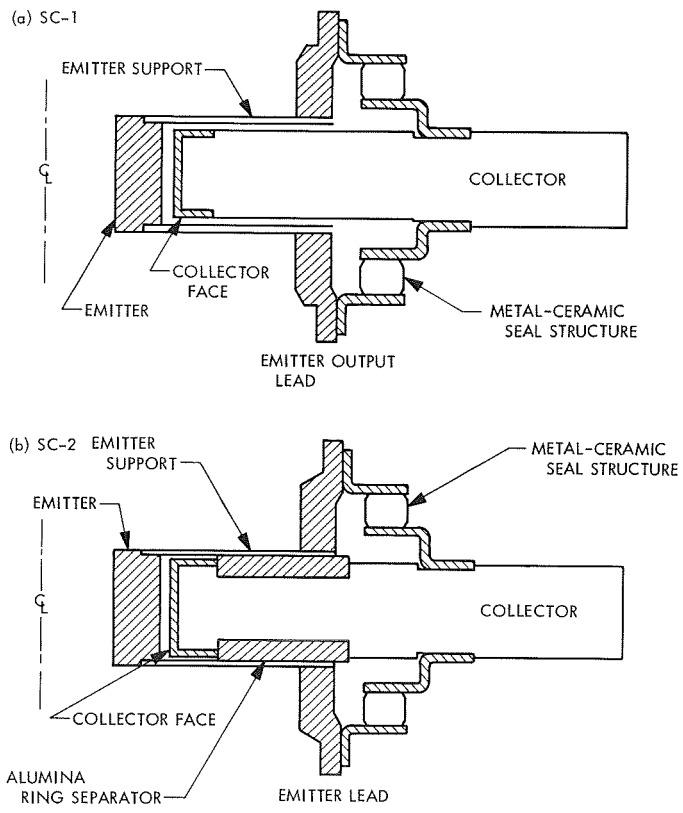


Fig. 22. SC-1 and SC-2 cross-sections

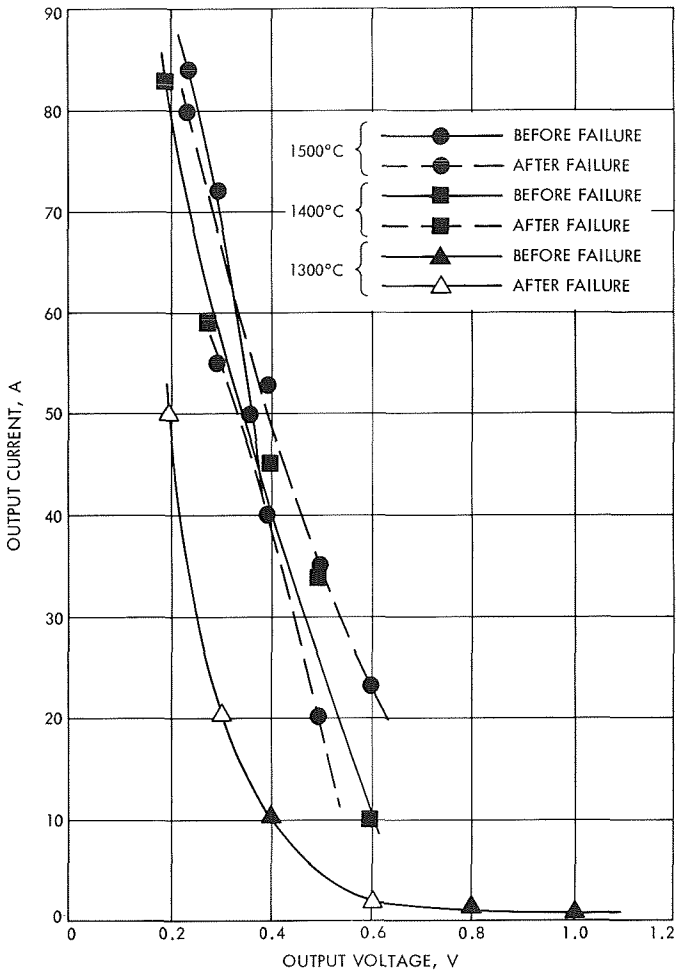


Fig. 23. SC-2 test results

tor lateral faces to prevent a short circuit as observed in SC-1. A schematic of the cross-section of SC-2 is presented in Fig. 22b.

The SC-2 was successfully tested at three values of emitter temperatures: 1300°C, 1400°C, and 1500°C over several values of voltage output. The results are graphically presented in Fig. 23. However, during the tests at the higher temperatures, the converter developed a cesium leak as a result of a failure in the emitter-to-emitter support membrane weld; this failure of the electron beam weld was attributed to thermal expansion. The converter was successfully repaired, reprocessed, and retested, the results being indicated as dotted lines in Fig. 23. The change in performance was tentatively attributed to a possible partial contamination of the electrodes during reprocessing and, in addition, to a possible lack of concentricity between the electrodes as a result of permanent deformation introduced by uneven thermal stresses.

It was also observed that the configuration of the electron gun utilized during the tests introduced stray bombardment, thereby overheating the emitter support membrane and increasing the thermal stresses.

Presently, a new and improved iteration of the converter (SC-3) is being assembled and will be tested with a modified electron gun to avoid the stray electron bombardment.

References

1. Long, J. R., "Output Power Characteristics of Thermionic Converter Array Under Partial Failure Conditions," *Advances in Energy Conversion*, pp. 163–170, Proceedings of the Intersociety Energy Conversion Conference, Miami Beach, Fla., Aug. 13–17, 1967.
2. Holland, T. W., "Performance of Cesium Thermionic Diodes Operated in Series-Parallel Circuits," Paper 2570-62 presented at the ARS Space Power System Conference, Santa Monica, Calif., Sept. 25–28, 1962.
3. Shimada, K., and de Winter, F., "Optimization of Thermionic Generator Systems of High Reliability," Proceedings of the Fifth Space Conference, p. 2.3.1-7, Cocoa Beach, Fla., Mar. 11–14, 1968.
4. Broido, T. H., *et al.*, *A Digital Computer Code for Analysis of Thermionic Networks*, Report GA-4147, General Atomic, Division of Gulf Co., May 22, 1963.

I. Testing of the SNAP-27 10-Couple Thermoelectric Modules, *G. Stapfer*

1. Introduction

The success of future space explorations will in part depend on the long-term performance of the electrical power source to be used for these missions. An investigation of the continuous performance of thermoelectric materials to be utilized for power generation is therefore not only desirable but necessary. As a result, life and endurance tests of the various proposed thermoelectric materials are presently being conducted at JPL.

Two lead–telluride, 10-couple thermoelectric modules were procured from the 3M Company for life-test evaluation. Both modules are similar in construction to the SNAP-27 thermoelectric generator presently being evaluated.

This article describes the life-tests performed on these modules, the equipment utilized for these tests, and the analysis of the resultant test data.

2. Description of Thermoelectric Module

Since both thermoelectric modules procured from the 3M Company are identical in design and content, a description of only one module is given.

The basic module contains the thermoelectric elements, an electrical heater, and a heat exchanger, all enclosed within an hermetically sealed canister. Figure 24 shows a pictorial view of the final assembly in the test station.

A total of 10 thermoelectric couples are connected in two series circuits of five couples each. The electrical leads for the two circuits are brought out separately through the canister on metal–ceramic feedthroughs. This arrangement allows the module to be operated in an all-series or series–parallel combination. For the purpose of the described life testing, both modules were operated with all 10 thermocouples wired in series. The temperature profile of the module is measured by Iron–Constantan thermocouples, which are attached to the hot and cold shoes.

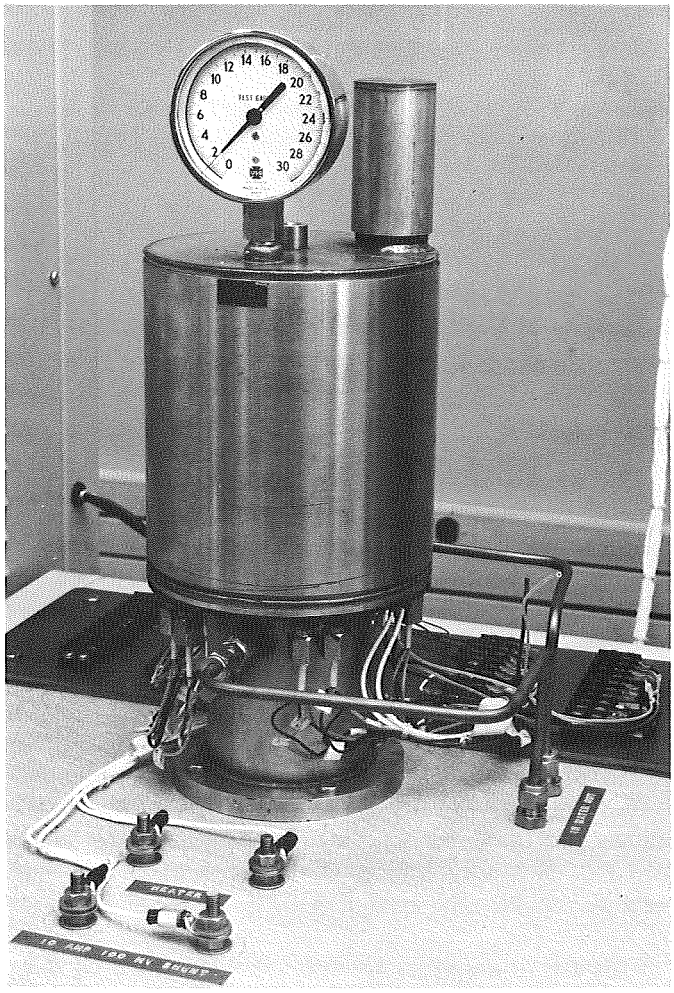


Fig. 24. 10-couple module

The thermoelectric elements are fabricated from 3P lead–tin–telluride material for the P-legs, and 3N lead–telluride for the N-legs. The thermoelectric materials are pressure contacted to the common hot shoe while the cold side is soldered to individual domed cold shoes. The dome of the cold shoe, through a spring follower, provides a heat path from the cold end of the couple to the module's cold frame. Excess heat is then removed from the cold frame by water cooling coils located between the cold frame and the canister walls.

The thermoelectric couples are thermally insulated from the housing by means of powdered MIN-K-1301. The same material is used for insulation between the hot and cold shoes. To remove any possible contaminants, the MIN-K-1301 is thoroughly outgassed prior to final assembly. Bottom and top plates are welded to the cylindrical assembly to obtain a hermetic enclosure. All the input and output connections for the heaters, instrumentation, etc., are brought through the canister bottom plate. The whole assembly is then purged and pressurized with argon to an internal pressure of approximately 2 psig at ambient temperature. This pressure is calculated to increase to about 8 psig during normal operating conditions. A Bourdon gage, mounted on the top plate, is used to continuously monitor the module internal pressure. A summary of the module specification is shown in Table 12.

Table 12. 10-couple module specification

Thermoelectric material:	
N-leg	Lead–telluride (3N)
P-leg	Lead–tin–telluride (3P)
Leg dimensions	0.207-in. diam 0.400-in. length
Operating temperature:	
Hot side	600°C nominal
Cold side	260°C nominal
Thermal insulation	MIN-K-1301
Fill gas	Argon

3. Module Test

a. Test equipment. The input power to the thermoelectric modules is controlled and maintained constant by a silicon-controlled rectifier power control circuit. A power input mode select switch maintains either the

power input or the hot junction temperature constant. The output of the power controller is rectified and filtered before being applied to the module heater. This not only reduces ac noise throughout the test system, but also increases the accuracy of the data to be recorded.

The output power of the module is continuously adjustable from zero voltage to open-circuit voltage by an electronic load. The module output voltage is maintained constant (± 10 mV) at the desired voltage level.

The temperatures of the module are monitored by a 10-channel thermocouple scanner. This temperature scanner also provides an over-temperature guard interlock, which protects the module from extreme temperatures by turning off all input power should any temperature exceed a predetermined value.

A digital voltmeter is used to measure all input and output voltages and currents and the module temperatures. The front panel meters are thus utilized only for visual reference and observations.

b. Life tests. Both modules were installed in the life-test equipment and the tests initiated in September 1966. Since that time, module 1 has accumulated over 19,000 h of operation and is still operating. Module 2 developed an open circuit after slightly more than 8000 h of operation. One-half of the module (five couples) is presently continuing the life test. In this article, however, only the data obtained prior to the open circuit are considered.

During the early test period, the internal pressure of both modules continuously decreased, indicating a loss of the hermetic seal. As a result, the modules were returned to the manufacturer for reprocessing and sealing. The subsequent life tests indicated a constant pressure for a portion of the operation; later, however, the pressure again began to decrease. The internal pressure of module 1 has decreased from an average of 8 psig to 2.5 psig, while module 2 pressure has decreased to atmospheric. No further attempt has been made to repurge either module.

Throughout the life test, numerous thermal cycles were experienced by the two modules. The thermal cycles were caused either by equipment failures or electrical power outages. In both cases, the modules were slowly brought back to the operating condition recorded prior to the shutdown. The module electrical power output was then compared with the previously recorded power output to verify the continuing performance. Any performance change directly attributable to the thermal cycling could thus be easily detected. However, at no time was such a performance change noted.

c. Test data. Throughout the major portion of the module life test, only limited attention was paid to test operation. As a result, data was recorded at intervals of approximately 50 h under varying conditions. Neither the input power to the modules nor the hot junction temperature was maintained constant. The module output voltage and current were allowed to vary over a large range of values.

The input and output power parameters of the two modules are shown in Figs. 25 and 26 as a function of operating time. The data for these curves were obtained

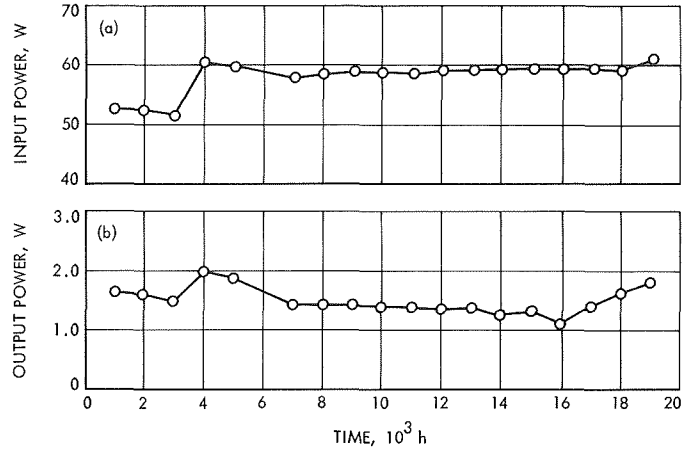


Fig. 25. Module 1 input and output power versus operating time

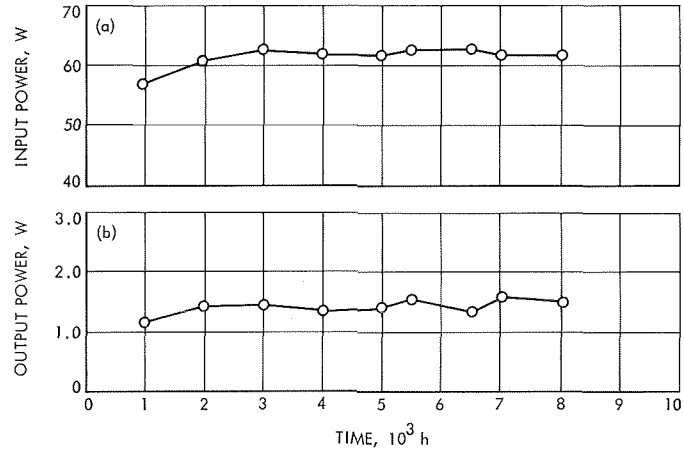


Fig. 26. Module 2 input and output power versus operating time

by selecting representative performance points from the recorded raw data at 1000-h intervals. The actual voltage and current measurements were used to derive the corresponding input and output power. The module's average hot junction temperature varied from 500°C to 600°C for module 1 and from 550°C to 600°C for module 2. A cold junction temperature of 250°C was recorded for module 1, while the average cold junction temperature of module 2 was 270°C. The open-circuit voltage was recorded automatically after the load circuit was interrupted (20–100 ms time delay). This ensured a fair degree of accuracy in recording this voltage since operator errors were eliminated.

4. Analysis of Data

To establish meaningful performance results from the module life tests, some normalizing of the recorded raw data was necessitated by the fact that throughout most of the module life testing, none of the input or output parameters were held constant. The initial normalizing consisted of adjusting the module output power as a function of input power. Since the module output power is proportional to the applied input power, a correction constant may be found and applied to this parameter. Data obtained from special tests showed that the power output will change by 0.031 W for each watt of input power change.

The module power output normalized to a constant input power of 62 W showed a smoother curve compared to the raw data; however, variations due to different operating voltage points still existed and distorted the actual output power profile.

To establish a true power degradation rate for these modules, it was necessary to determine the output power at a constant output voltage. Since it is advantageous to use a reference voltage that will correspond to the maximum power output point of the module, the value of 0.8 V was selected. From a recent volt–ampere curve recorded with a constant input power of 62 W, the ratio of the power at a specified output voltage to the power at the reference voltage (0.8 V) was established. The corresponding ratio is then applied to the normalized output power of the module. The resultant voltage-normalized power output P_{vn} thus represents the power output of the module had it been operated at the reference output voltage with a constant input power of 62 W.

Figure 27 shows the P_{vn} of the two modules that have been normalized to a constant voltage and input power

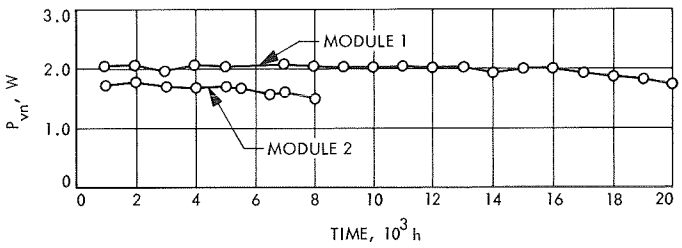


Fig. 27. Modules 1 and 2 voltage-normalized output power versus operating time (reference output voltage = 0.8 V)

as a function of operating time. It can be seen from this figure that the output power of module 1 decreased very little for the initial 16,000 h of operation; however, a noticeable increase in the degradation rate is apparent beyond this point.

The initial power degradation rate of module 1 is 0.1 W for the first 16,000 h or 0.3%/1000 h. This rate subsequently increased to 0.25 W/4000 h or 3.12%/1000 h. The tenfold increase in the power degradation of the module may be caused by the small leak in the hermetic seal (the pressure decreased to about 4 psig at the 16,000-h operating point). Since the internal gas pressure is 4–6 psi lower in the module's cold (shutdown) state, it is possible that air was introduced into the module during an inadvertent shutdown. The resulting presence of oxygen would seriously affect the subsequent performance of the module.

The power output of module 2 is also continuously decreasing. A change of degradation rate became apparent after approximately 5500 h of operation. The power degradation rate up to this point is 0.45%/1000 h; subsequently, the rate increased to 6%/1000 h.

The degradation rates of both modules are in fair agreement, and it appears that the same process is responsible for the change in the degradation rate of module 2 as was discussed for module 1. To establish whether the observed output power degradation is due to either a change in the Seebeck voltage or internal resistance, both of these parameters need to be further examined.

Making use of the relationship

$$E_{oc} \simeq \alpha (T_H - T_C)$$

where E_{oc} is the open-circuit voltage, α is the Seebeck coefficient in μV/°C, T_H is the hot junction temperature, and T_C is the cold junction temperature, the open-circuit

voltage can be normalized to an average hot junction temperature. Using a value of 5mV/°C as a temperature correction factor, it was found that the open-circuit voltage remained fairly constant at 1.45 V throughout the tests. This indicates that the change of internal resistance is solely responsible for the output power degradation.

The internal resistance of the module is composed of the actual thermoelectric material resistance in addition to the interconnecting contact resistance. It is very difficult to determine which of the two resistances changed during operation. However, since the contact arrangement uses spring tension, it is possible that this tension changes as a function of operating time, and thus contributes to the overall internal resistance increase.

5. Conclusions

The two 10-couple thermoelectric modules have successfully operated for a combined total of 27,000 h and their overall performance has been found to be very stable. The observed output power degradation, inherent with thermoelectric power sources, was found to be due to an increase in the internal resistance of the module. With the available test data, it is difficult, if not impossible, to determine whether the observed increase of internal resistance is due to a deterioration of the contacts or a change in the physical properties of the thermoelectric material.

The change in the power degradation rate of module 1 is possibly due to the introduction of air into the module during inadvertent shutdowns. This becomes a possibility when the normal operating pressure falls below 4–6 psig. It was found, however, that this change of degradation rate will be completely masked unless the test parameters are held extremely constant or some normalizing of the raw data is performed.

J. The Development of a Long-Life, High Cycle Life, 30-A-h, Sealed AgO–Zn Battery,

R. E. Patterson and R. S. Bogner

1. Introduction

Some aspects of the AgO–Zn system make it desirable for space application usage. Among these aspects are its high energy density (4/1 weight advantage over Ni–Cd), non-magnetic nature (no residual magnetic field), low capacity loss on room temperature stand (less than 1%/mo), and clear end-of-charge indicator (sharp voltage rise). However, future missions contemplated by NASA/JPL impose stand-life and cycle-life requirements that cannot be met by current model AgO–Zn batteries. In order to capitalize on the system's inherent advantages, JPL is engaged in development activities for improving the stand life and cycle life of the AgO–Zn system.

Specifically, ESB, Inc. is presently under contract to JPL to develop a 30 A-h high cycle life, sealed AgO–Zn battery. The major design goal is for 100 or more charge/discharge cycles at 50% depth of discharge after a 6-mo activated stand operating in a temperature range of 50–100°F. Currently, 12 different design groups of 5 cells each have been built and are under test (Phase I). In addition, five design groups of 22 cells each are now being fabricated and will be placed under test (Phase II). Design variables include type of separator, electrolyte concentration, active material ratio, and negative electrode contour.

This article discusses the motivation for the Phase I and Phase II experimental designs and the test plan of each.

2. Cell Designs Selected for Phase I (Ref. 1)

The experimental design variables for 12 different 30 A-h cell designs (5 cells per design) are presented in Table 13 and include two electrolytes, two ZnO:Ag ratios, five separator systems, and three wrapping styles.

Two concentrations were selected for test. These are 45 and 41% KOH, 90% saturated with ZnO. For long life, the upper concentration is desirable to retard Ag ion diffusion from positives to negatives and oxidative hydrolysis of cellulosic membranes. The lower concentration gives better charge and discharge voltages and ampere–hour efficiencies at the C/2 rate at 50°F. The effect of the concentration change is expected to be greater for separator systems having proportionately larger fractions of irradiated polyethylene membranes.

J. Lander (Ref. 2) has observed that the cycle life of 24 A-h, sealed AgO–Zn cells increases from 300 to 1140 cycles on a 2-h orbit at 25% depth of discharge and 75°F in 45% KOH when the molar ratio of Zn:Ag is increased from 1:1 to 2:1. Based on this and other data molar ratios, 1.59:1 and 1.86:1 were selected. Both ratios are greater than that of *Mariner* Mars 1969 cells (1:1).

Three wrapping styles to be tested are (1) positive wrap, (2) negative wrap, and (3) combined wrap. For long-term

Table 13. Phase I experimental design

Contour of negative plate	Electrolyte concentration, % KOH (90% saturated with ZnO)	ZnO: Ag weight ratio	Separator system											
			Positive wrap				Negative wrap						Combined	
			1 layer EM-470 4 layers FSC		1 layer EM-470 4 layers RAI-2291 2 layers FSC			1 layer EM-476I 6 layers RAI-2291			1 layer EM-470 4 layers RAI-2291 2 layers FSC		3 layers RAI-2291[a] 1 layer EM-476I 4 layers RAI-2291[b]	
			Design 2	Design 2′	Design 3	Design 3′	Design 1	Design 1′	Design 4	Design 4′	Design 5	Design 5′	Design 6	Design 6′
Wedge	45	1.2:1					5							
Standard	45	1.2:1	5[c]		5			5			5			
Standard	41	1.2:1		5		5			5			5	5	
Standard	41	1.4:1								5				5

[a]Positive electrode.

[b]Negative electrode.

[c]Refers to number of cells in the group.

Table 14. Common design features

Active Ag density	69 g/in.^{3}
Positive process	Sintered
Negative active mix	
ZnO	90%
HgO	3%
Teflon	7%
Negative process	Sintered
Negative plate density	45 g/in.^{3}
Negative grid	2/0 distorted
Pack tightness	1.25 psi (minimum)
Cell jar material	ABS
Cell jar sealant	Epoxy and ABS cement
Platelock	Epocast 221/927
Zn moss block	Turned over top edge of separator system

cycling on a 22-h charge and 2-h discharge, the cell reaction and osmosis will favor transfer of water from positive to negative compartments, thus raising electrolyte levels around negatives, lowering levels around positives, which will tend to starve positives and polarize the cell. The small pore diameter of RAI-2291 should aggravate this condition. On the basis of this hypothesis, capacity should be maintained best for the system with least fluctuation of levels and least concentration changes at positives and negatives. A predicted ranking applicable for requirements of this development activity in order of effectiveness would be:

(1) Negative wrap.

(2) Positive wrap.

(3) Combined wrap.

One cell of each test group will be manufactured in a clear cell jar to permit observation of levels during cycling.

Test separators will be two absorbers (irradiated Kendal EM-476 and Kendal EM-470) and two membranes (Visking fibrous sausage casing (FSC) and RAI-2291). The combination of separator systems in the cell designs of Table 13 is intended (1) to determine the system most capable of retarding Ag penetration, Zn penetration, and negative plate errosion; and (2) to improve both cycle and activated life over the present *Mariner* Mars 1969 cell system.

Design features common to all designs are presented in Table 14.

3. Cell Designs Selected for Phase II

Designs 2 through 6 (Table 13) were selected for Phase II from the designs of Phase I with the ZnO:Ag molar ratio and the KOH concentration fixed at 2:1 and 43%, respectively. Design constants listed in Table 14 are applicable to the Phase II designs.

The intent of Phase II is to study the effectiveness of the five different separator systems at the battery level. This can be done since 22 cells of each design will be fabricated. The effect of design variables of Phase I will be observed at the cell level.

4. Phase I and Phase II Test Plan

Cells of Phase I and Phase II will be cycled through 100 or more cycles (Table 15). Cycling will require a modified constant potential 22-h charge and a constant resistance 2-h discharge. Cycling will begin with charge after a discharge simulating a typical encounter. When the five-cell groups of Phase I and the 19-cell groups of Phase II (Table 15) no longer deliver required capacity to 1.0 V on the lowest cell, that cell will be removed and cycling continued. Four 100% depth capacity measuring cycles will be performed during cycling approximately 25 days apart as a measure of negative plate errosion. Following cycling, representative cells from each group will be disassembled to determine failure modes.

5. Anticipated Results

Since the Phase I cells have just been placed on test and the Phase II cells are only now being activated, it is too

Table 15. Test plan for Phase I and Phase II

Event	Cells per group																			
	Phase I												Phase II							
Four cycles and recharge	5	5	5	5	5	5	5	5	5	5	5	5	19	3	19	3	19	3	19	3	19	3
Vibration														3		3		3		3		3
4-mo charged stand at 70 ±3°F	5	5	5	5	5	5	5	5	5	5	5	5										
6-mo charged stand at 70 ±3°F													19	3	19	3	19	3	19	3	19	3
Fifth cycle discharge and recharge	5	5	5	5	5	5	5	5	5	5	5	5	19	3	19	3	19	3	19	3	19	3
90 or more autocycles, 50% depth-of-discharge, 70°F, plus four 100% depth-of-discharge, capacity measuring cycles	5	5	5	5	5	5	5	5	5	5	5	5	19	3	19	3	19	3	19	3	19	3
Disassembly and inspection	2	2	1	1	1	1	2	2	1	1	1	1	5	1	5	1	5	1	5	1	5	1
Design number	1	1′	2	2′	3	3′	4	4′	5	5′	6	6′	2		3		4		5		6	

early to draw any definite conclusions. However, when testing is completed, it is expected that test results will give some indication as to:

(1) The suitability of AgO–Zn batteries for use on missions requiring battery cycling after a transit period of up to 6 mo.

(2) The effect of KOH concentration on capacity loss, stand life, and cycle life.

(3) The effect of the ZnO:Ag ratio on capacity loss, stand life, and cycle life.

(4) The effect of a contoured negative plate on capacity loss, stand life, and cycle life.

(5) The effectiveness of the different combinations of separator systems in retarding Ag–Zn penetration and errosion, and improving both cycle and activated life.

(6) The applicability of extrapolating cell test data to predict battery performance.

References

1. Farris, C. D., First Quarterly Report on JPL Contract 952472, ESB Report E-23-69. ESB Inc., Exide Missile and Electronics Division, Raleigh, N.C., June 16, 1969.

2. Lander, J., *Sealed Ag–Zn Batteries*, Paper 64-749, presented at the AIAA Third Biennial Aerospace Power Systems Conference Sept. 1–4, 1964.

K. RTG Radiation Test Laboratory, *R. W. Campbell*

1. Introduction

The need and overall objectives for a radioisotope thermoelectric generator (RTG) test laboratory have been discussed in SPS 37-52, Vol. III, pp. 46–47. It was pointed out that the future requirement for solar-independent power sources for spacecraft has led to the use of RTGs and to JPL's role in developing the technology required to ensure their integration on advanced mission spacecraft. In support of this effort, it was indicated that work in two major areas was required: (1) life and performance testing of both fueled and electrically heated RTGs, and (2) determination of the compatibility of the radioisotope heat source, and its inherent gamma and neutron radiation fields, with space science instrumentation and other sensitive subsystems. To fulfill the requirements in this latter area, JPL has now developed an RTG Radiation Test Laboratory and has embarked on an experimental program with the following test objectives: (1) the determination of both energy and intensity of the neutron and gamma-ray radiation emitted by RTGs, (2) determination of the level of interference from this mixed radiation field with scientific instruments and spacecraft subsystems and the deliniation of how this interference varies according to radiation component and energy of that component, (3) determination of methods for reducing the interference with sensitive components, (4) performance of long-term radiation effects measurements.

This article will describe the work that has gone into making this laboratory operable, features and capabilities of the laboratory to serve as background for future articles on work coming out of the laboratory, and some of the present endeavors.

2. Facility

In October 1968, JPL received final approval from NASA to modify Building 247 for use as the RTG Radiation Test Laboratory. A contractor was chosen and construction was begun the last week of January 1969. By April 1, 1969, modifications were completed to the point where occupancy was possible and laboratory set-up could begin.

The laboratory building is underground and located in a low-population density area on Arroyo Road near the east border of JPL. It is surrounded by a 5-ft high fence to limit access to the area. The overall dimensions of the building are 67 × 16 × 9 ft with the roof approximately 5 ft below ground level. The walls, floor, and ceiling are reinforced concrete approximately 14-in. thick. Access to the building from ground level is by an entry way external to the main building. The interior is divided by a 2-ft thick reinforced concrete wall into a small storage area (13 × 16 ft) and a large experimental area (54 × 16 ft).

The building design includes the radiation safety requirements for the use of radioisotope heat sources and other radioactive materials. By adding absolute filters to an existing exhaust system, it was possible to design the facility so as to contain any release of radioactive airborne contamination within the laboratory building itself.

Since the building may contain classified material, it has the design features necessary to classify it as a strong room.

During the approval and construction phases of the facility, a standard operating procedure for the laboratory

was developed and written to satisfy the requirement to be included by the AEC for the receipt of a *Systems for Nuclear Auxiliary Power* (SNAP)-27 radioisotope heat source. All major items of this document have been approved by the AEC Albuquerque operations office. In addition, the JPL state license for the possession of radioactive material was amended to allow acquisition of radioactive sources necessary for carrying out an adequate experimental program.

3. Equipment

An extensive amount of equipment was procured during this construction phase to equip the laboratory for both the planned experimental programs and for an adequate radiation safety monitoring program.

a. Radiation safety. The major items of instrumentation procured for radiation monitoring are:

(1) AIM-3 airborne alpha particle detector.

(2) Radeco airborne plutonium alpha particle detector.

(3) PAC-4C scintillation alpha counter.

(4) Planchett proportional counter system.

(5) Victoreen Model 440 survey meter.

(6) Ray-d-Tec radiation meter (beta, gamma).

This instrumentation was checked out and calibrated by the JPL instrument services group and is in current operation. The instrumentation is divided between that used daily within the lab and that located in an emergency cabinet external to the lab to ensure the availability of uncontaminated instruments in the unlikely event of an extensive radioactive spill. Also located in this emergency cabinet are self-contained breathing apparatus hoods, coveralls, boots, and gloves for use if reentry into such an area would be required prior to decontamination.

b. Experimental. For the experimental measurement of gamma and/or neutron spectra, five major spectrometer systems have been assembled. They are:

(1) 20-cm^3 Ge (Li) gamma-ray spectrometer.

(2) 3 × 3-in. NaI (Tl) gamma-ray spectrometer.

(3) NE-213 liquid organic scintillation neutron spectrometer.

(4) He^3 solid-state sandwich neutron spectrometer.

(5) Thermoluminescence detector system.

c. Radiation sources. Gamma-ray calibration sources were obtained that were calibrated in absolute intensity traceable to NBS standards. On arrival, all sources were leak-checked to ensure their integrity. The sources and the major radiation can be seen in Table 16. Since RTG fuel capsules will not be available at all times, a variety of other radiation sources has been acquired. A 5-Ci PuBe neutron source has been purchased. Eight Pm^{147} heat sources of approximately 3000 Ci each have been obtained from a source within JPL. Six SNAP-15A Pu^{238} metal heat sources of ≈1.5 W each have been received on loan from Mound Laboratory through the AEC. In addition, a simulated PuO_2 source has been designed that will accurately reproduce the gamma-ray spectrum expected from a SNAP-27 fuel capsule without the use of high-activity alpha emitters. The source also makes it possible to mock-up the changes in the radiation spectrum expected in a PuO_2 fuel capsule over an 18-yr period. This source is discussed in *Section F* of this chapter.

4. Experimental Program

a. Gamma-ray spectrometer. A series of measurements was made with the 20-cm^3 Ge (Li) solid-state spectrometer system, using the calibrated gamma-ray sources, in order to determine its absolute counting efficiency. Seven calibration sources were used covering an energy range from 0.166 to 2.60 MeV. The gamma-ray emission was known to an accuracy of between 2 and 8%.

The data in the photopeaks were analyzed by least-square fits to a Gaussian plus a straight-line function. This technique (generally credited to Gauss) linearizes the

Table 16. Standard gamma-ray calibration sources

Radioisotope	Energy of major γ-ray, keV
Am^{241}	59.5
Cd^{109}	87.7
Hg^{203}	279.2
Cr^{51}	320.1
Sr^{85}	514.0
Cs^{137}	661.6
Mn^{54}	834.8
Zn^{65}	1115.4
Co^{60}	1173.2 and 1332.2
Na^{22}	1274.5
Ba^{133}	Many energys
Co^{56}	Many energys

function by expanding it in a Taylor series with respect to each of the parameters and truncating the series after the linear terms. This was done using a computer code called Gauss III (obtained from Phillips Petroleum, Idaho). When a log–log plot of efficiency versus energy was made, the data were found to lie along a straight line. A weighted least-square fit of the data of $\ln E$ versus $\ln \epsilon A$ was made, where E is the energy in MeV, ϵ is the absolute efficiency for a gamma ray that strikes the detector giving a pulse that falls in the full energy peak, and A is the frontal area of the detector. The weights were calculated by combining the calibration error and the error in the photopeak area. The final result was $\epsilon A = 0.383 E^{-1.330}$. The error in ϵA may be calculated from the variances and covariances in the parameters. At 1 MeV, the error in the efficiency is 1.7%. (The area of the detector is arbitrary as it will cancel out when the efficiency is applied to the measurement of an unknown sample.)

b. NE-213 organic scintillation neutron spectrometer. This system was assembled and the electronics adjusted to give proper discrimination against gamma-ray pulses. This was done using a PuBe neutron source plus additional gamma-ray background sources. Since the laboratory has no source of monoenergetic neutrons at present, the spectrometer system was then sent off lab for calibration (see *paragraph d*).

c. SNAP-15A Pu^{238} metal source. The gamma-ray spectrum emitted by a SNAP-15A Pu^{238} metal source was investigated in detail using a 20 cm^3 Ge (Li) detector. The source is doubly encapsulated in a right-circular cylinder about ¾ in. long by about ⅜ in. in diameter. The source-to-detector distance was about 1 ft. Various portions of the spectrum were viewed on a 400-channel pulse-height analyzer used in conjunction with a biased amplifier. The area under the photopeaks is presently being analyzed in order to make an absolute intensity determination. To do this, the photopeak efficiency versus energy curve for the crystal obtained during its calibration is used.

d. Radiation interference measurements. A large number of nuclear detectors that form the components of the most sensitive scientific flight instrument packages have been acquired. These detectors will be used for radiation interference measurements. The detectors include:

(1) EON-6213 Geiger–Muller counter.

(2) Bendix 4010 channeltron.

(3) Silicon surface barrier detector (fully depleted).

(4) Photomultiplier tubes (several types).

(5) Pilot B plastic scintillator.

(6) Cesium iodide (Na) crystals.

The response of the Geiger–Muller counter, silicon surface barrier detector and channeltron are now under investigation using the SNAP-15A Pu^{238} metal heat source as the radiation source. Measurements are also being made of the efficiency versus energy using the calibrated gamma-ray sources.

e. Battelle Northwest Laboratories contract. A contract was awarded to Battelle Northwest Laboratories, Richland, Washington, to determine experimentally the absolute efficiency of certain scientific flight instrument components as a function of incident monoenergetic neutron and photon radiation both with and without shielding materials. Also included are some irradiations to study radiation damage. The initiation date of the contract was May 19, 1969.

The first phase of the program now being carried out includes the procurement of the detectors and the design and fabrication of the shields required. The types of detectors that have been purchased are (1) miniature Geiger–Muller counters, (2) lithium drifted silicon detectors, (3) totally depleted silicon detectors, and (4) electron multipliers.

In addition to the measurements of the efficiency of these scientific flight instruments, Battelle will also perform absolute calibrations of one or more of the neutron spectrometer systems to be used at JPL to determine the neutron spectrum from SNAP heat sources. It is presently calibrating the NE-213 detector system.

5. Interfaces with Other Facilities

a. TRW. TRW has been contracted by the NASA–Ames Research Center, Moffett Field, California, to perform a study to ascertain the extent of possible RTG radiation interference on science instruments for deep-space probes and, in particular, those instruments selected for the *Pioneers F* and *G* missions. JPL has negotiated, through Ames Research Center, to receive portions of the raw experimental data obtained from this contract. At JPL, these data will be analyzed so as to make them directly comparable to experimental information obtained in our own program as well as to our parallel analytic effort, both of which are considerably broader in scope than contracted for in the TRW work.

Arrangements were also made for JPL to make a limited number of gamma-spectra measurements of the SNAP-27 fuel capsule assembly at TRW using the laboratory's calibrated Ge (Li) solid-state gamma-ray spectrometer and 3×3-in. NaI crystal. These measurements are presently being analyzed.

b. Goddard Space Flight Center. Communication with personnel in the RTG effort at Goddard Space Flight Center has revealed that both facilities can benefit considerably by an exchange of computer programs developed for the reduction of experimental data. Such program developments represent a significant expenditure in time and money and any duplication of effort in this area should be avoided. JPL has already received program CUBED for the analysis of a continum spectrum with a limited number of superimposed peaks obtained from NaI (Tl) spectrometer systems. In its present form, the code is somewhat limited for JPL's purposes due to the small number of peaks that can be treated; however, it is presently being upgraded at the Goddard Space Flight Center and the newer version (CUPED) will be sent to us shortly. The JPL RTG Radiation Test Laboratory has awarded a contract to NUS Corporation to further modify this improved version to treat data obtained from the Ge(Li) system, which will then be made available to the Goddard Space Flight Center.

XII. Spacecraft Control

GUIDANCE AND CONTROL DIVISION

A. Inertial System Tests, *G. Paine*

1. Introduction

The SEAN system is a strapdown navigation system employing three digital velocimeters (DVM) and two electrically suspended gyros (ESG) for inertial references. This system has been described previously (SPS 37-52, Vol. III, pp. 52–55, and SPS 37-56, Vol. III, pp. 133–135).

This report describes results of tests on the system in the laboratory with only one of the gyros operating. The tests were conducted with the sensors stationary. The digital flight computer program was the complete two-gyro program, augmented by a driver subroutine which simulated the missing gyro. The driver simulation uses optical measurements of the inertial measurement unit's (IMU) orientation to compute the appropriate gyro outputs. Thus, although nonstationary tests cannot be performed using this technique, stationary tests can be, and they provide a useful vehicle for continued checkout of the flight computer program and a good insight into overall system behavior.

The results of one such test are shown in Fig. 1 for a favorable gyro spin axis orientation. In this test no gyro drift compensation is used.

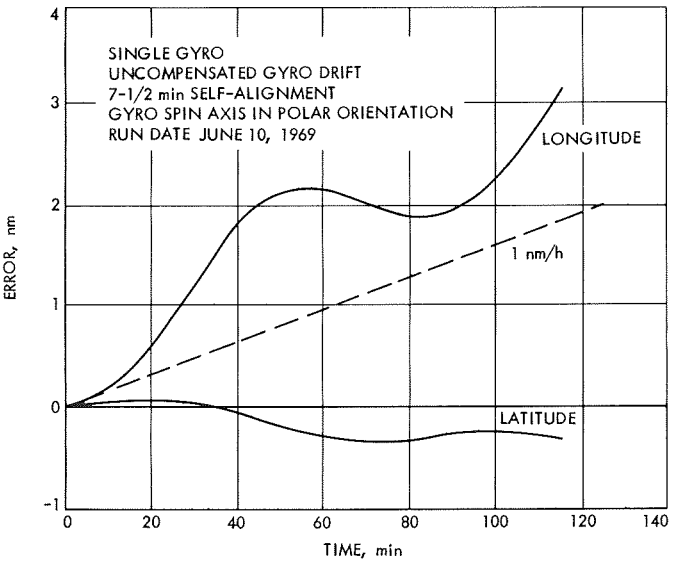


Fig. 1. SEAN system performance

2. Test Conditions

These tests were performed in a laboratory designed for the test of inertial components. The laboratory is surveyed to first order and consequently optical measurements with respect to a North, East, vertical coordinate system can readily be obtained. Throughout these tests

the IMU was mounted on a Goertz air-bearing servo-table. The table position was fixed during each test. This constraint being necessitated by the lack of two gyros.

Attached to the IMU is a cube which defines the IMU (body) reference system. The misalignments of both the DVM's and the gyros are referenced to this cube.

The DVM's had been calibrated 3 mo prior to the tests. Since these tests were not the final system performance tests, more recent measurements were not deemed necessary. The single gyro used was calibrated for pickoff misalignments and rotor pattern shortly before the tests began. Sets of coefficients for the prediction of gyro drift were also available for certain special gyro orientations.

The system, with the exception of one gyro and the altimeter which will be installed prior to van testing, was complete. That is, in a form suitable for performing self-contained navigation in a van or aircraft environment.

3. Software

The software employed was principally mechanized for the Honeywell Alert flight computer. One additional program was written for the Univac 1108 and used to edit the data stored on magnetic tape during a test by the digital recording system (DRS). The Alert computer used in SEAN has a 24-bit word, a 2-μs memory cycle time, and 8192-word memory, and a direct memory access channel for the DRS.

The various major subroutines, their functions, core requirements, and execution times are shown in Table 1. The total storage requirements are thus about 7500 words, not including a 576-word buffer used by the DRS. In the worst cycles, one complete iteration takes about 51 ms. Worst case refers to both instruction execution times and computational load, since not all subroutines are computed on any particular cycle, although they may be.

Four features of this program deserve special mention: the driver, the error equations, the editing subroutine, and the DRS. The first and second of these both make use of the optical measurement of the IMU cube. This cube is used to define the body coordinate system of the SEAN system (Fig. 2). With respect to this cube both down and cross range errors are defined. The driver takes the optical measurements and computes the inertial orientation of the cube, defined at system turn-on. From the inertial orientation it computes the "counts" that a perfect gyro would sense. The error equations compare the driver's computed orientation with that from the navigation program to arrive at a figure of performance for the SEAN system. The editing routines used in this series of tests were not capable of handling the noise, and

Table 1. SEAN Alert program subroutines

Subroutine	Function	Core, word	Time,[a] ms
	Prealignment sequence	800	
	Input routine	300	0.3
D	Driver	400	4.9
ED	Gyro edit	200	
EROR	Calculate alignment errors	200	
TRBG	Calculate body–gyro transformation matrix	500	9.9
VTRT	Vehicle turning rate	300	9.1
CTAC	Compute centrifugal and tangential acceleration compensation	100	~5.0
GYDR	Gyro drift	600	7.2
ALGN	Alignment	1000	
NAVI	Navigation integration, compensations, gravity computations	600	11.9
BLIP	Display	300	2.2
TAPE	DRS dumps	100	~0.8
DMUL	Double-precision subroutine	100	
SQ	Square root, sine/cosine, arc sine subroutine	200	
MATS	Double-precision matrix routines	200	
	Working storage	600	
	DRS buffer	600	
	Keyboard, paper tape input	1000	
		8100	51.3

[a]Times are included for a navigation cycle only. Computation time spent in utility routines, e.g., square root, sine, etc., are included in calling routine time.

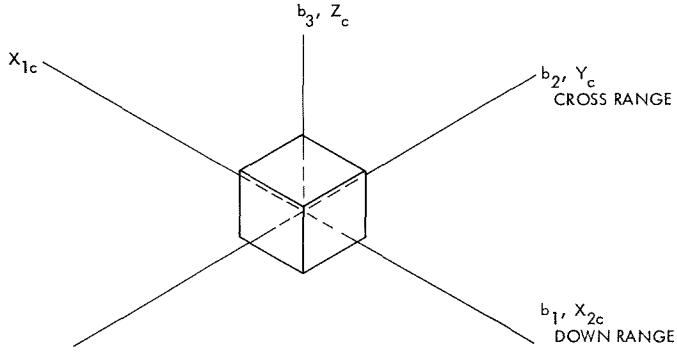


Fig. 2. IMU cube body coordinates

system failures associated with the ESG outputs in all tests and are being revised. The gyro rotor contained a few imperfections on its surface and on the pattern which disturbed the optical readout of the pickoff–rotor spin axis direction cosine. Finally the DRS recorded on magnetic tape the data contained in a 576-word buffer area about fifteen times per minute. This buffer was loaded at the same frequency with the most important variables.

The Univac 1108 DRS program used to process the Alert DRS magnetic tapes performs the task of converting the numbers from octal to decimal, selecting some parameters for special display, and reducing the data to be examined to a manageable level. The program was also capable of handling a memory dump mode in which the entire contents of core would be dumped onto magnetic tape.

4. Results

Only the results from one particular test are shown here. In this particular test the ESG spin axis was in a polar orientation. This is an orientation that results in low drift and only small amounts of bad data that would have to be edited. The gyro drift was not compensated. The self-contained alignment was programmed for about 7.5 min. The IMU cube was oriented with two axes nearly level and nearly pointing East and North. Thus, down-range errors are in longitude, and cross-range errors are in latitude. Figure 1 gives a time history of the errors in this test.

B. Digital Systems for Inertial Sensors, *P. J. Hand*

The all-digital feasibility study system and single-axis simulator for the *Thermoelectric Outer Planet Spacecraft* (TOPS) program requires a digital gyro system that will provide serial digital position information to the attitude-control system (ACS).

The electronic design is a pulse-on-demand system, based in a large part on past successful JPL experience in both analog and digital circuits. The ±20-V precision regulator is identical to the *Mariner* Mars 1971 unit; the ac amplifier demodulator–dc amplifier loop is similar to the *Mariner* Mars 1971 design, except for gains; and the comparator used is taken from its accelerometer system. Choice of component parts is largely from the present Preferred Parts Catalog. The basic mechanization of this pulse-on-demand system is similar to that of the *Mariner* Mars 1971 accelerometer loop, except that it is bi-directional rather than uni-directional. The loop has a limited dynamic range as far as capture capability is concerned and only modest linearity over the working range.

Gyro drift performance is substantially unaffected by the digital loop and remains the same as for an analog system. Resolution of the first form of the loop was limited to 0.005 deg/pulse with a capture capability of 1.0 deg/s. This resolution was 1/10 the pointing accuracy requirement of the single-axis simulator and was considered to be sufficient. It has since been shown that this is not adequate for autopilot steering requirements, and even finer resolution will be required of the final gyro loop in a TOPS system.

The gyro selected for use with this breadboard digital system is the Honeywell, Inc., type GG334S. Several working units were available at the conclusion of the sterilizable wide-angle gas-bearing gyro program. This GG334S gyro design is suitable for digital torquing, in that the torquer scale factor is 350 deg/h/mA. The gyro gain is 0.4, which is also compatible with pulse torquing requirements.

The GG334S is a single-degree-of-freedom floated rate-integrating, hydrodynamic gas-bearing motor gyroscope. The all-ceramic gimbal is supported by ball pivots and a dithered jewel system. The motor, except for the magnetic parts, is also made from ceramic. Rotor speed is 24,000 rpm, producing an angular momentum H of 100,000 gm-cm^2/s. Motor driving power is 800-Hz two-phase, 26-V rms at approximately 3.3 W. The operating temperature is 180°F, providing a damping C of 2.5×10^5 dyne-cm-s. Thus a nominal gyro gain (H/C) of 0.4 is obtained. The GG334S is a wide-angle gyro, in that an input angular freedom of ±12 deg is provided. This fact is of little importance for use as a digital gyro in a torque rebalance loop as the gimbal is always constrained within a few arc seconds of null. A block diagram of the first form of the rebalance loop is shown in Fig. 3.

The gyro pickoff is excited by 9.6 Khz at 5.0 V. This signal is derived from a digital divider chain and is referenced to the 38.4-Khz crystal oscillator. After filtering the 9.6-Khz square wave to a sine wave, it is amplified and power-boosted to drive the pickoff primary. This same 9.6-Khz square wave is also the transistor demodulator reference signal. AC gain of the pickoff preamplifier is 15.0, while the combined gain of the demodulator and dc amplifier is 67.5. The total gimbal position signal channel gain is 1013, giving a scale factor of 100 V/deg of input axis motion. This gimbal position signal is applied

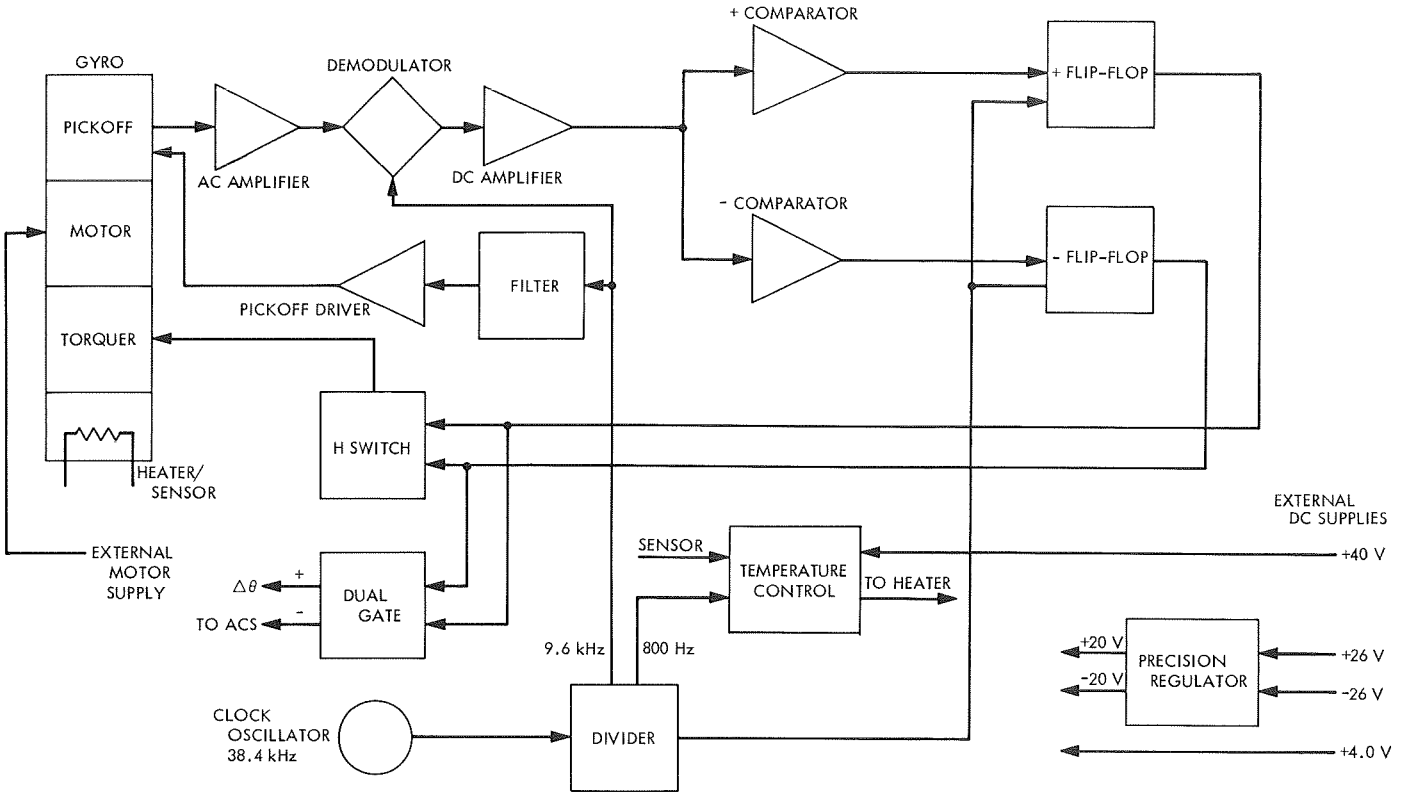


Fig. 3. Rebalance loop

to dual voltage comparators, whose trigger levels are set at 0.5 V. One comparator changes from a logical zero to a logical one for a positive input level, the other changes similarly for a negative input voltage level.

Whichever comparator is triggered enables a control flip-flop, which is also driven by a 400-Hz bi-phase clock signal. The enabled control flip-flop will produce one complete pulse for each two transitions of the clock signal, thus dividing the clock signal in half. The enabled control flip-flop is coupled to the appropriate two quarters of an H switch, which also contains level shifting and inverting functions. The H switch output is a precision current pulse of 22 mA through a current limiting resistor (1800 Ω for 0.005 deg/pulse) to the gyro torquer. This current, which is on for 5.000 ms, is just sufficient to return the gimbal position signal to null at low input rates. The control flip-flop outputs are also buffered by a dual gate and sent to the ACS as a logical *one* signal indicating a $\Delta\theta$. Two separate $\Delta\theta$ lines are used, one for each polarity.

Gyro temperature control is provided by a digital method utilizing high-efficiency pulse-width modulation techniques. In spite of the fact that the pulse-width modulator is switching 40 V at 425 mA in about 1 μs or less, no significant noise problems have been noted in the first version breadboard.

Spin motor power for the breadboard is provided by an external 800-Hz 2-phase sine wave supply; also, four external Hewlett Packard 723A dc supplies are used for the electronic functions. The entire breadboard is built on 7 Vector Co. plugboard cards.

Initial calibration of the pulse weight ($\Delta\theta$) to within $\pm 0.2\%$ was made with the gyro on an optical dividing head (Leitz) in the input axis vertical orientation. The entire system breadboard was then moved to the Fecker Type 152 servo-table for dynamic measurements of the $\Delta\theta$ value, as well as for gyro drift tests. These tests were run over a 10-day period. The test method was first to allow the gyro to sense vertical earth rate (8.45497 deg/h at the inertial sensor laboratory) over about a 30-min period after warm-up and prior to the start of the dynamic run and repeated after the run was completed. At least three sets of 100 period averages were taken for each drift measurement. The difference between indicated

earth's rate from the gyro and the actual rate is listed in Table 2 as the combined drift error. It is the combined effect of spin-axis mass unbalance, fixed torques, and reaction torques on the gyro's ability to sense 8.45 deg/h. These effects cannot be separated without changing the gyro's orientation on the table; however, past analog testing has shown the largest contributor from the standpoint of instability, is the spin motor reaction torque term. Shifts in its value of 0.3 deg/h have been observed during conventional analog testing. All drift effects were then trimmed out by injecting a small current through a very large resistor (1.8 MΩ) into the torquer. The servotable was next commanded to turn at the required rate for the test. The Fecker Microgon table position readout system was used to gate a totalizing counter for a segment of 100 deg. This counter's other input was the $\Delta\theta$ output from the gyro loop. The Microgon has a basic accuracy of approximately ±3.6 arc seconds on a static basis. The test setup would, therefore, display the total number of rebalance pulses generated during each 100 deg of input motion. Anywhere from 3 to 57 sets of 100-deg readings were taken in one continuous run. Both positive and negative outputs were recorded at various times, as well as several different input rates from 650 deg/h to 0.9 deg/s. The results of these runs are also shown in Table 2. The only significant change in day-to-day stability was shown on the last day, with a large shift in drift values and a decrease in $\Delta\theta$ of about 0.8%.

Table 2. Drift and calibration data—GG334S gyro S/N B1 digital breadboard (first version)

DATE	Drift		Input rate		Calibration deg/pulse
	Indicated earth's rate, deg/h	Combined error, deg/h	pulse/s	deg/s	
9/5/69	8.1113	−0.3437	+35.4	0.177	0.005 004 50
			+99.1	0.496	0.005 004 25
			−100.3	0.502	0.005 004 50
			−179.4	0.897	0.005 004 75
			+179.5	0.897	0.005 004 50
9/9/69	8.0419	−0.4131	+98.4	0.492	0.005 004 75
	7.9177	−0.5373	+180.4	0.902	0.005 004 75
9/12/69	8.0343	−0.4207	+98.4	0.492	0.005 004 25
9/15/69	8.5300	+0.0750	−36.1	0.181	0.005 001 25
	8.3288	−0.1262	+36.3	0.182	0.005 000 75
			+97.7	0.489	0.005 000 75
			−98.5	0.493	0.005 001 25
			+179.6	0.898	0.005 000 75
			−179.5	0.493	0.005 001 50

The first breadboard form of the digital gyro loop has shown adequate stability, sufficiently good linearity, and a minimum number of parts to demonstrate the feasibility of this approach.

C. New Methods for Evaluating the Performance of Approximate Nonlinear Filters, *A. K. Bejczy*[1]

1. Introduction

There are guidance and control systems warranting the application of nonlinear filtering for the purpose of effective control. The analysis and synthesis of nonlinear filters, however, are associated with a wide variety of mathematical and structural complexities. These complexities are often undesirable because of hardware or other constraints. Thus, when dealing with nonlinear filtering problems the primary concern is the computational complexity of the filter algorithms (Ref. 1).

Any nonlinear filter algorithm is inherently approximative. This is due to the fact that, in general, a nongaussian conditional probability density cannot be represented by a finite number of elementary functions. For practical reasons, therefore, approximations are inescapable.

A number of different approximate nonlinear filter schemes have been derived in the last few years. (A short account on the different approximate nonlinear filters can be found, e.g., in Ref. 2). The hitherto known nonlinear filter schemes, however, have an important structural feature in common: the approximate statistics of the filtering error is described by stochastic differential equations which are coupled to the stochastic differential equations governing the time behavior of the estimated mean of the state of a dynamical system. (It is worthwhile to recall that this structural feature forms a sharp contrast to the structure of the linear Kalman filter where the statistics of the filtering error is described by deterministic differential equations which are decoupled from the stochastic differential equations governing the time behavior of the estimated mean of the state of a

[1]Research partially completed during tenure of a National Research Council postdoctoral resident research associateship at JPL, supported by NASA.

dynamical system.) Thus, the known approximate nonlinear filter algorithms are entirely of stochastic character. This fact has far-reaching consequences regarding the performance evaluation as well as the implementation of the filter algorithms. The only known method of determining the performance of the approximate nonlinear filters is by Monte Carlo type numerical studies. Those studies are usually very time consuming and, hence, expensive. Moreover, the implementation of a filter algorithm of an entirely stochastic nature may often be unfeasible.

A new formulation of the approximate nonlinear filtering problem is currently under development in a joint effort together with R. Sridhar[2]. The full work will be soon published in detail in a joint paper elsewhere. In the present article only the headlights of the new theory will be briefly described and, some numerical examples illustrating the revenues of the new method will be disclosed.

2. Theoretical Aspects of the New Method

In the new formulation of the approximate (linear or nonlinear) filtering problem the functional form of the filter gain can be postulated. Then, the new formulation adjoins the equation of the filter to the original process equation after eliminating the observation variable from the filter equation. Introducing suitable assumptions, it can be shown that the enlarged vector is a Markov process which satisfies a Langevin equation. The transition density of this process is governed by the Fokker–Planck equation. The filtering error is the difference between certain components of the enlarged vector. The density function of the filtering error can then be obtained as marginal density from the density function of the enlarged vector which satisfies the Fokker–Planck equation. Approximate equations for the mean and for the covariance of the filtering error can be readily derived. These are ordinary (deterministic) differential equations which can be directly solved, at least numerically. Thus, in the new formulation of the approximate filtering problem the performance of the approximate filters can be evaluated by deterministic techniques, bypassing the Monte Carlo type of studies. At the same time, sufficiently simple approximate filter configurations can be constructed by the new method. This is a very important result from the point of view of feasible implementation.

[2]With the Department of Electrical Engineering, California Institute of Technology.

In the framework of the new formulation of the approximate filtering problem there are different levels of approximations which can be introduced to deriving approximate equations for the mean and for the covariance of the filtering error. The resulting simplest and most desirable approximate equations for the mean μ and for the covariance β of the filtering error are

$$\dot{\mu} = [f'(t,\mu) - G(\mu)h'(t,\mu)]\mu \tag{1}$$

$$\dot{\beta} = [f'(t,\mu) - G(\mu)h'(t,\mu)]\beta + \beta[f'(t,\mu) - G(\mu)h'(t,\mu)]^T + G(\mu)QG(\mu)^T + R \tag{2}$$

where

μ = an n-vector

β = an $n \cdot n$ symmetric matrix and

$f(t,\cdot)$ = system equations, n-vector function

$h(t,\cdot)$ = observation equations; m-vector function; $m \leq n$

$G(\cdot)$ = filter gain; $n \cdot m$ matrix

Q = error covariance matrix for the measurement noise which is assumed to be gaussian

R = error covariance matrix for the system noise which is assumed to be gaussian, too, and uncorrelated to the measurement noise

In Eqs. (1) and (2) the dot denotes time derivative, the prime denotes derivatives with respect to the state, and T denotes the transpose of a matrix. Furthermore, if $\mu(0) = 0$ then, according to Eq. (1), $\mu(t) \equiv 0$. Consequently, Eq. (2) has to be evaluated at $\mu = 0$ which results that Eq. (2) will be expressed in terms of known system and statistical parameters.

Eqs. (1) to (2) are belonging to the following simple approximate nonlinear filter scheme:

$$\dot{\hat{x}} = f(t,\hat{x}) + G(\hat{x})\,[y(t) - h(t,\hat{x})] \tag{3}$$

where $y(t)$ is the running observation, and the hat ($\wedge$) signifies the estimated value of the state vector x. In Eq. (3) G may also be taken as a constant.

3. Numerical Results

The new method is currently tested on several nonlinear filtering problems. The results are very promising. As an illustration for the revenues of the new method the following scalar case will be quoted here.

The nonlinear system is given by

$$\dot{x} = -\frac{x}{1+x^2} + n_S \tag{4}$$

Two observation equations are selected:

$$\begin{aligned} y &= x + n_M \qquad \text{linear} \\ &= \arctan(x) + n_M \qquad \text{nonlinear} \end{aligned} \tag{5}$$

The initial condition is given in terms of a normal distribution $N(\cdot,\cdot)$ with specified mean and standard deviation, respectively:

$$x(0) \; : \; N(\alpha,\sigma_0) \tag{6a}$$

The system and measurement noise are also given in terms of a normal distribution:

$$n_S \; : \; N(0,\sigma_S) \tag{6b}$$

$$n_M \; : \; N(0,\sigma_M) \tag{6c}$$

a. Results of the new method. For convenience, a constant filter gain is postulated, denoted by κ. According to Eq. (3) the approximate nonlinear filter equation for the estimated mean of the state is then

$$\dot{\hat{x}} = -\frac{\hat{x}}{1+\hat{x}^2} + \kappa[y - \hat{x}] \qquad \text{linear} \tag{7a}$$

or

$$\dot{\hat{x}} = -\frac{\hat{x}}{1+\hat{x}^2} + \kappa[y - \arctan(\hat{x})] \qquad \text{nonlinear} \tag{7b}$$

The initial condition for Eqs. (7a) and (7b) is $\hat{x}(0) = \alpha$.

According to Eq. (2) and, assuming $\mu(0) = 0$, the approximate variance of the filtering error is given by the following very simple first-order ordinary linear differential equation:

$$\dot{\beta} = -2(1+\kappa)\beta + \kappa^2\sigma_M^2 + \sigma_S^2 \tag{8}$$

In the present problem Eq. (8) is valid both for the linear and the nonlinear measurements. The initial condition for Eq. (8) is $\beta(0) = \sigma_0^2$.

Eq. (8) admits a solution in closed form:

$$\beta(t) = \frac{\sigma_S^2 + \kappa^2\sigma_M^2}{2(1+\kappa)} + \left(\sigma_0^2 - \frac{\sigma_S^2 + \kappa^2\sigma_M^2}{2(1+\kappa)}\right)\exp[-2(1+\kappa)t] \tag{9a}$$

with a steady-state ($t \to \infty$) value

$$\bar{\beta} = \frac{\sigma_S^2 + \kappa^2\sigma_M^2}{2(1+\kappa)} \tag{9b}$$

The value of the constant filter gain κ that minimizes the steady-state variance $\bar{\beta}$ will be

$$\kappa^* = \left[1 + \left(\frac{\sigma_S}{\sigma_M}\right)^2\right]^{1/2} - 1 \tag{10a}$$

yielding the following minimum value for the steady-state variance of the filtering error:

$$\bar{\beta}^* = \sigma_M^2\,\kappa^* \tag{10b}$$

It has to be emphasized that Eq. (8) (or alternatively, Eq. 9a) describes the time history of the variance of the filtering error in terms of known quantities in a simple, deterministic way. In the framework of the new method it is also a simple matter to predict the steady-state value of the variance and to determine the value of the constant filter gain that will minimize the steady-state value of the variance. This value of the constant gain can be termed the optimum constant gain, while the corresponding filter can be called the optimum constant gain approximate nonlinear filter.

The reliability of the new method for evaluating the performance of the approximate nonlinear filters has been tested by Monte Carlo simulations (MCS). The procedure of the MCS for the problem of interest is schematically described in Fig. 4.

The MCS for the problem of interest contains three major parts:

(1) A random number generator producing three pools of gaussian random numbers with specified mean and standard deviation for the initial conditions on the original process equation, Eq. (4); for the system noise n_S; and for the measurement noise n_M.

(2) A computer integration routine (Runge–Kutta–Gill algorithm) acting on the original process equation, Eq. (4), and on the filter equation, Eq. (7a) or (7b).

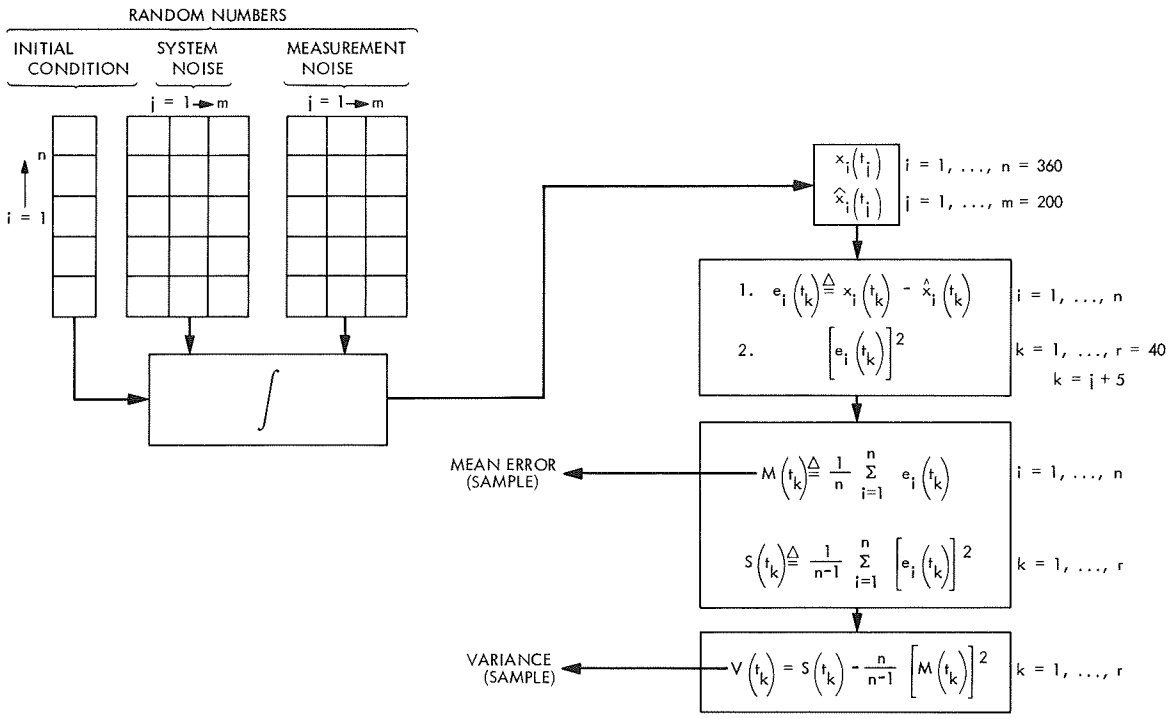


Fig. 4. Schematics of Monte Carlo simulation of approximate nonlinear filters

(3) Computations of the sample mean and of the sample variance at selected time points.

In the MCS the sampling time is made equal to the integration step size. That is, at each integration step Δt a new relevant random number is fed into the integration routine for the system noise and for the measurement noise, respectively. These random numbers are held constant during that integration step.

In generating random numbers with specified standard deviation, allowance is made to the fact that the computer solution of the differential equations is in the discrete time domain, while the formulation of the filtering problem is posed in the continuous time domain. The specified continuous random processes n_S and n_M are translated into equivalent discrete random sequences $\bar{n}_S$ and $\bar{n}_M$ in the random number generator through the following equivalence claim:

$$\left.\begin{aligned} \bar{\sigma}_M &= \frac{\sigma_M}{(\Delta t)^{1/2}} \\ \bar{\sigma}_S &= \frac{\sigma_S}{(\Delta t)^{1/2}} \end{aligned}\right\} \tag{11}$$

where Δt is the integration step size.

In the performed MCS studies the total integration interval was 20 s with an integration step size $\Delta t = 0.1$ s. (For the example under consideration shorter step size did not change the results in any significant manner.) The sample mean and the sample variance are computed only at each fifth step, that is, in 0.5-s intervals. The sample space contained 360 integration runs for each MCS study. This provides a high confidence value for the obtained MCS results. Some of the obtained MCS results together with the belonging results of the new theoretical performance evaluation method are depicted in Figs. 5 and 6. These figures show clearly that the sample variance does indeed converge to the theoretically predicted (approximate) variance of the filtering error.

Figure 5 displays two cases: (a) system noise to measurement noise ratio $\sigma_S/\sigma_M = 2$; (b) $\sigma_S/\sigma_M = 0.5$. In both cases, however, the constant filter gain is the optimum one, computed according to Eq. (10a). As it is seen in Fig. 5, the predicted (theoretical) minimum steady-state variance $\bar{\beta}^*$ is identical with the sample (experimental) minimum steady-state variance. Minor differences between the theoretical and experimental values of the transient variance, however, are to be expected—as it is seen in Fig. 5—because of the simplifying assumptions involved in deriving the approximate theoretical equations, Eqs. (1) and (2).

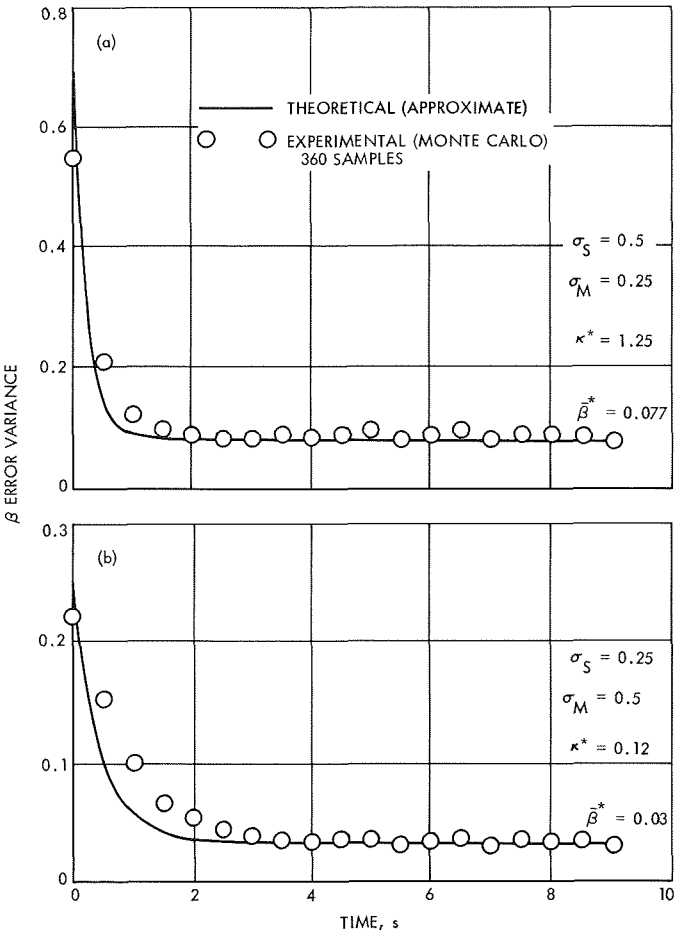


Fig. 5. Performance of optimum constant gain nonlinear filters

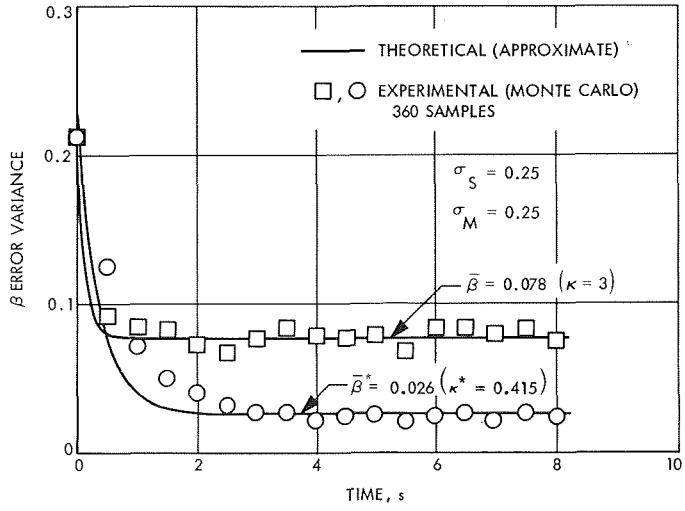


Fig. 6. Performance of optimum and of non-optimum constant gain nonlinear filters

Figure 6 displays a case with $\sigma_S/\sigma_M = 1$ and with two different values for the constant filter gain: the optimum one and another gain higher than the optimum one. As it is seen in Fig. 6, the optimum constant gain really yields the minimum steady-state variance; the gain higher than the optimum one, however, tends to shorten the transient time for the variance. It is also seen in Fig. 6 that the theoretical steady-state variance (minimum or nonminimum) is identical with the experimental steady-state variance in each case.

b. Results of other approximate methods. It is worthwhile and very instructive to compare the new performance evaluation method (Eqs. 7 to 8) to the other known approximate nonlinear filter algorithms applied to the same filtering problem (Eqs. 4 to 6). For convenience, only the maximum principle least-squares (MPLS) nonlinear filter-algorithm (Ref. 3) is selected here for the purpose of comparison.

The MPLS nonlinear filter-algorithm as applied to the filtering problem given by Eqs. (4) to (6) yields the following coupled system of nonlinear stochastic differential equations:

Linear measurements

$$\dot{\hat{x}} = -\frac{\hat{x}}{1+\hat{x}^2} + \frac{P}{\sigma_M^2}[y - \hat{x}] \tag{12}$$

$$\dot{P} = 2P\frac{\hat{x}^2 - 1}{(1+\hat{x}^2)^2} - \frac{P^2}{\sigma_M^2} + \sigma_S^2 \tag{13}$$

for the estimated mean and variance, respectively.

Nonlinear measurements

$$\dot{\hat{x}} = -\frac{\hat{x}}{1+\hat{x}^2} + \frac{P}{\sigma_M^2(1+\hat{x}^2)}[y - \arctan(\hat{x})] \tag{14}$$

$$\dot{P} = 2P\frac{\hat{x}^2 - 1}{(1+\hat{x}^2)^2} - \frac{P^2}{\sigma_M^2(1+\hat{x}^2)^2} - \frac{2\hat{x}P^2}{2\sigma_M^2(1+\hat{x}^2)^2}[y - \arctan(\hat{x})] + \sigma_S^2 \tag{15}$$

for the estimated mean and variance, respectively.

The initial conditions are $\hat{x}(0) = \alpha$, $P(0) = \sigma_0^2$ in both cases.

Eqs. (12) and (14) should be compared to Eqs. (7a) and (7b), while Eqs. (13) and (15) should be compared to Eq. (8) in order to illustrate the features of the new

performance evaluation method. In the new method Eq. (8) (which is a measure for the filter's performance) can be solved *independently* of Eqs. (7a) and (7b). While in the MPLS algorithm Eqs. (13) and (15) (which represent the measures for the filter's performance) must be solved *together* with Eqs. (12) and (14), respectively. The solution of Eq. (8) is a *deterministic* solution, while the solutions of Eqs. (13) and (15) must be performed in a *probabilistic* sense, that is, by Monte Carlo methods. From the point of view of implementation, furthermore, the MPLS algorithm calls for the mechanization of Eqs. (12) and (13) or of Eqs. (14) and (15), while the new method calls for the mechanization of Eq. (7a) or of Eq. (7b) alone, which obviously is a much simpler task than the former one.

To illustrate the revenues of the new method in more depth, Eqs. (12) and (13) and Eqs. (14) and (15) were solved by MCS method as it was described in the previous subsection. (See also Fig. 4). The obtained results for a case with $\sigma_S/\sigma_M = 1$ are depicted in Fig. 7.[3] For the sake of comparison, the result of the new method (the solution of Eq. 8 using the relevant optimum constant gain) is also depicted in Fig. 7. As it is seen in Fig. 7, the MCS solution of Eq. (13) (or, of Eq. 15) is very close to the solution of Eq. (8). It is also important to note that both the P and the κ^* filters yield the same experimental (sample) variance. Consequently, the performance of the rather complex P filter is, for the problem under consideration, not better in any practical manner than the performance of the simple κ^* filter constructed by the new approximation method.

[3]Note that the present case is different from the case depicted in Fig. 6, since now σ_S and σ_M are twice as high as they were in the computation depicted in Fig. 6.

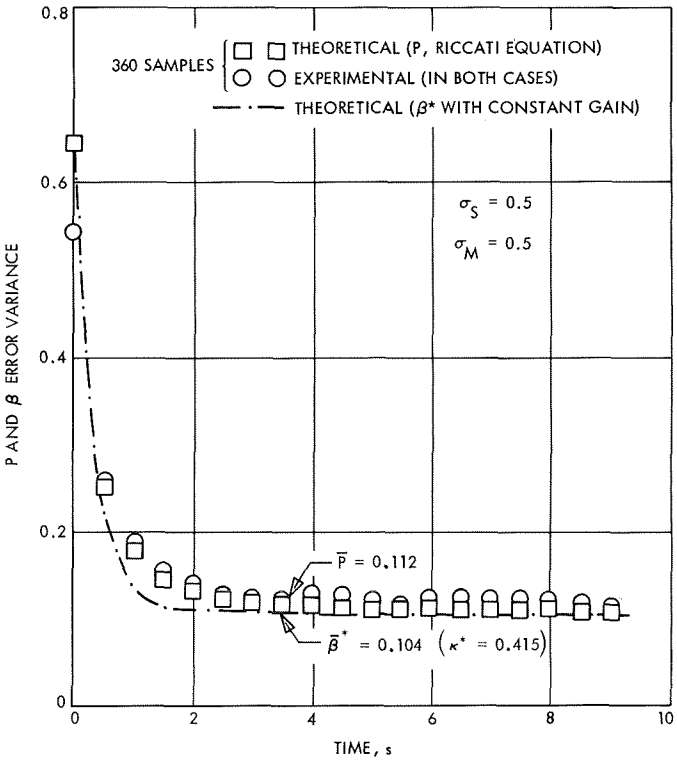


Fig. 7. Performance of optimum constant gain and of the MPLS nonlinear filters

References

1. Bejczy, A. K., *Sequential Estimation of States of a Ballistic Vehicle in an Imperfectly Known Atmosphere,* Technical Report 32-1405, Jet Propulsion Laboratory, Pasadena, Calif., October 1969.
2. Schwartz, L., and Stear, E. B., "A Computational Comparison of Several Nonlinear Filters," *IEEE Trans. Auto. Cont.*, pp. 83–86, February 1968.
3. Detchmendy, D. M., and Sridhar, R., "Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical Systems," *J. Basic. Eng.*, Vol. 88, Series D, No. 2, pp. 362–369, 1966.

D. Design and Prototype Development of the Speed Sensor and Optimal Controller for a Minimum Energy Controller, *Y. E. Sahinkaya*

1. Speed Sensor

The primary function of the speed sensor is to measure the speed of the dc motor and utilize this information in the establishment of steady-state operation for the minimum-energy controller SPS 37-56, Vol. III, pp.138–144). The motor speed is measured by a tacho-generator which is connected directly to the motor shaft. The output signal from tacho-generator is fed to the input circuit of pulse-width modulator (SPS 37-57, Vol. III, pp. 99–103) i.e., to the base circuit of transistor denoted by Q_5, after passing through a transistor amplifier. The output signal of this transistor amplifier adjusts the pulse width according to the steady-state form of the optimal control law during the motoring and generating modes of operation.

The two additional functions of the speed-sensor are as follows:

First, it supplies a signal to the armature-current direction sensor in order to determine the mode of operation. The output signal from the tacho-generator is fed to the input circuit of armature-current direction sensor, i.e., to

the base circuit of transistor denoted by Q_1, after passing through a second transistor amplifier. The output signal is proportional to the back emf voltage of the motor. The signal whose value is governed by the optimal control law is also fed into the input circuit of armature-current direction sensor, i.e., to the base circuit of transistor denoted by Q_2, after passing through a transistor amplifier. The resulting difference voltage between these two signals is then used to determine the mode of operation according to the principles given in SPS 37-57, Vol. III, pp. 99–103.

Second, it supplies a signal to the optimal controller. The output signal from the tacho-generator is fed to the input circuit of the optimal controller, i.e., to the base circuit of transistor denoted by T_1, after passing through a third transistor amplifier. The output signal of this transistor amplifier is compared with the output signal from the throttle-setting circuit, and the resulting difference voltage is then utilized to turn on or turn off the voltage supply for the acceleration or deceleration control circuits of the optimal controller.

The tacho-generator is placed in a balanced bridge circuit. The tacho-generator is driven mechanically by the drive motor. A signal, which is linearly proportional to motor speed, is applied to the base circuit of transistor denoted by Q_9 in Fig. 8. The transistors denoted by Q_9, Q_{10} and Q_{11} transfer this speed information to pulse-width modulator, armature-current direction sensor and optimal controller respectively, as shown in Fig. 8. The noise produced by the brush-commutator action of the tacho-generator is minimized by designing R-C filters shown in Fig. 8. Figure 9 shows the mathematical model which is utilized in the design of filters.

From Fig. 9

$$\frac{1}{C}\int_0^t i_2(t)dt + R_1 i_1(t) = v_{in}(t) \tag{1}$$

$$\frac{1}{C}\int_0^t i_2(t)dt = (R_2 + R_3)i_3(t) \tag{2}$$

$$i_1(t) = i_2(t) + i_3(t) \tag{3}$$

$$R_3 i_3(t) = v_{out}(t) \tag{4}$$

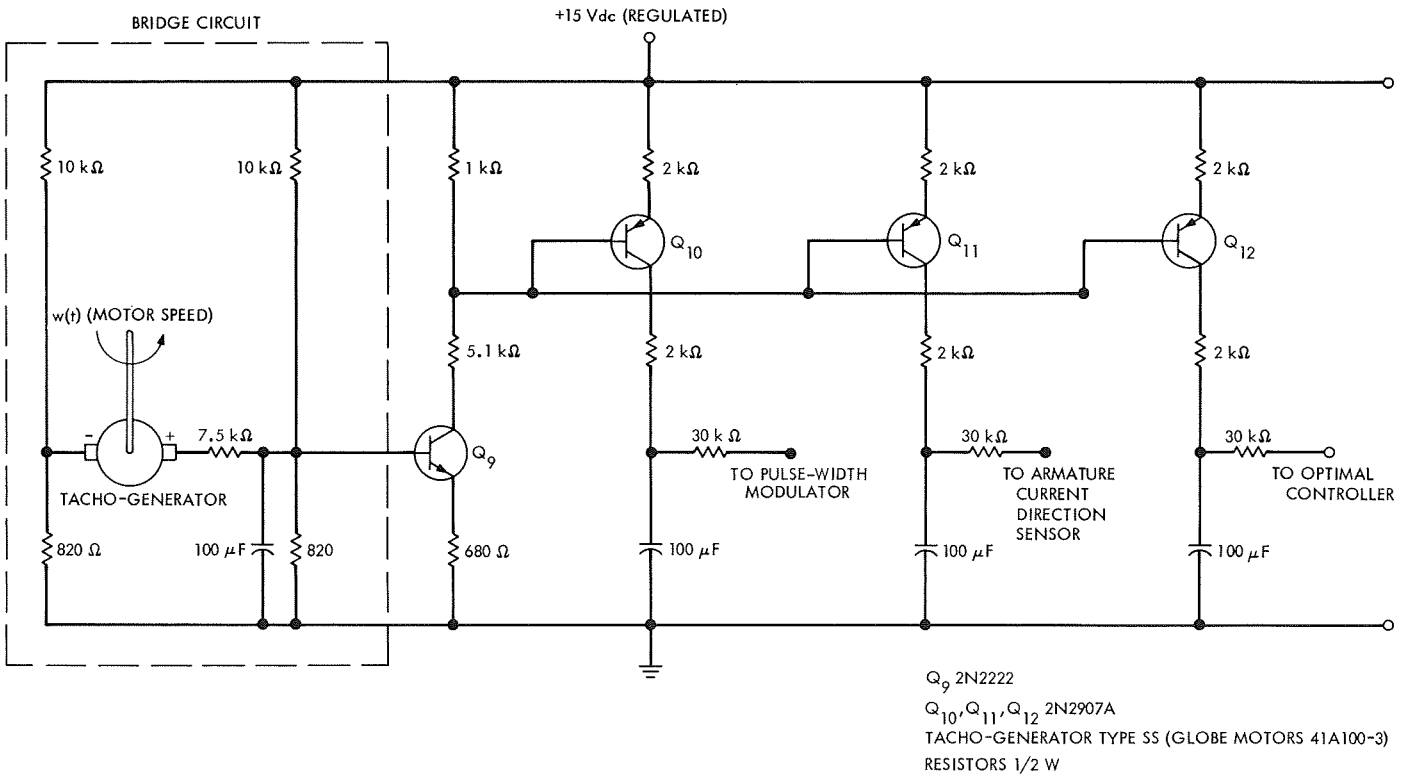


Fig. 8. Speed sensor schematic

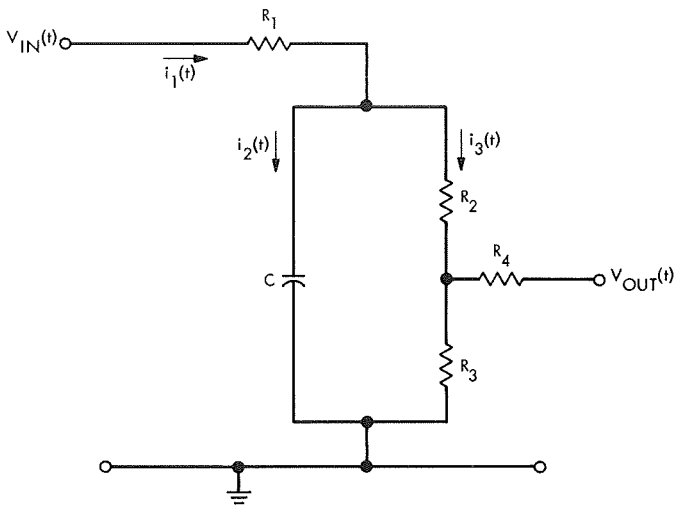


Fig. 9. Mathematical model for speed sensor filter

where

$$R_4 \gg R_3$$

Let

$$v_{in}(t) = v_{dc}(t) + v_{ac}(t) \tag{5}$$

where

$v_{dc}(t) \triangleq v_0$ = an arbitrary constant value of the dc level in the output signal of tacho-generator

$v_{ac}(t) \triangleq v_n \sin \omega_n t$ = undesirable noise signal with an arbitrary constant amplitude v_n and frequency ω_n in the output signal of tacho-generator.

The definitions of v_{dc} and v_{ac} indicate that only the steady-state operation of the motor is considered.

By utilizing Eqs. (1) through (5) the dc value of $i_2(t)$ is obtained:

$$i_2(t) = \left(\frac{v_{out}}{R_1}\right) \exp\left[-\left(\frac{1}{CR}\right)t\right] \tag{6}$$

where

$$R \triangleq \frac{R_1 R_4}{R_1 + R_4}$$

and the ac value is

$$i_2(t) = \left(\frac{v_n}{R_1}\right) \frac{\omega_n x}{(1 + \omega_n^2 x^2)^{1/2}} \cos(\omega_n t - \tan^{-1}\omega_n x) \tag{7}$$

where

$$x \triangleq RC$$

In Eq. (7), as ωx increases from zero to infinity,

$$0 \leq \frac{\omega_n x}{(1 + \omega_n^2 x^2)^{1/2}} \leq 1 \tag{8}$$

for all possible combinations of ω and x.

And if

$$\omega_n x = 1 \tag{9}$$

$$\frac{\omega_n x}{(1 + \omega_n^2 x^2)^{1/2}} = 0.707 \tag{10}$$

If a reasonable ripple frequency ω_n corresponding to a reasonably low speed of operation is assumed, then using Eq. (9) yields a design value for RC time constant of the filter.

Hence

$$i_2(t) \approx \left(\frac{v_n}{R_1}\right) \cos(\omega t - 45 \text{ deg}) \tag{11}$$

Equations (6) and (9) are used to determine the filter parameters shown in Fig. 8 by taking into account the following considerations:

(1) A small RC value in comparison with the response time of the minimum-energy controller is desirable. This is seen from Eq. (6).

(3) The chosen value of RC should satisfy the criterion for $\omega_n x$ as given by Eq. (9) at a realistically chosen value of ω_n. The filtering operation becomes even better for ω_n's in the higher side of the spectrum with respect to the design value of ω_n.

2. Optimal Controller

The primary function of the optimal controller is to respond to signals from the throttle-setting mechanism,

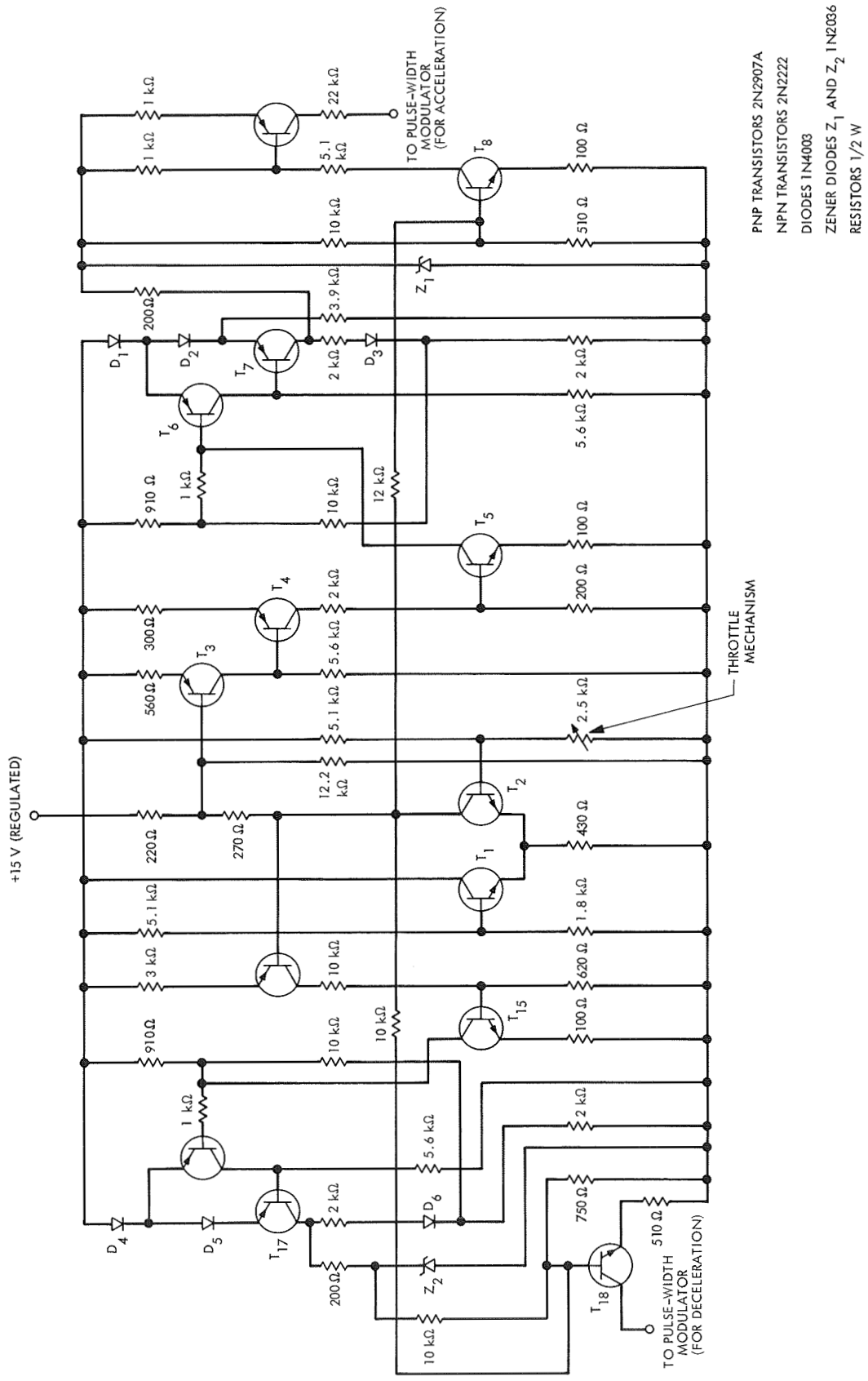


Fig. 10. Optimal controller schematic

i.e., desired speed, tacho-generator speed sensor, i.e., actual speed, and to accelerate or decelerate the motor (or vehicle) speed to the desired speed setting in an optimal manner. The throttle-setting mechanism also automatically sets the desired response time for the minimum-energy controller. The output signals from the optimal controller for acceleration or deceleration control of the vehicle are applied to the base circuit of transistor Q_5 in the pulse-width modulator. Whenever the desired speed is reached, i.e., $\omega(t) = \alpha$, as explained in SPS 37-56, Vol. III, pp. 138–144, the output signals of optimal controller becomes identically zero until a new desired speed setting is applied to the system. Design of the optimal controller has been accomplished utilizing standard analytical tools. Figure 10 shows the final values of the circuit elements of the optimal controller.

3. Test Results

The test results on the speed sensor and optimal controller prove to be satisfactory and agree very closely with their predicted behaviors. The testing of the integrated control system is still in progress.

E. Strapdown Electrically Suspended Gyroscope Development, *T. J. Donlin*

The effort on the strapdown electrically suspended gyroscope (ESG) has completed the current development stage. Progress on the ESG development was last reported in SPS 37-55, Vol. III, pp. 130–131.

The purpose in developing the ESG was to obtain a long-life accurate inertial reference for spacecraft attitude control. The results from this development period have established the level of ESG pointing accuracy that can be expected. The results have also indicated what is needed to refine the ESG error compensation.

During this ESG development effort two strapdown gyros were developed and built. Associated with the gyro effort, an ESG mathematical model was hypothesized and computer programmed to allow prediction of the gyro drift error behavior and the assigning of all the sources of the error within the gyro itself. Prediction coefficients in the mathematical model were assigned to particular sources of error, such as mass unbalance torques; primary electric torques; second, third, and fourth-order electric torques; and translation torques. A symmetrical positive definite nonsingular matrix has been set by using a least-squares method. Eigenvalue and eigenvector solutions were used to examine the conditioning of the matrix. One drift test alone could not give enough information about the ESG to separate the torque coefficients. A simulation study was conducted to determine the direction in which the gyro calibration test development should proceed.

As stated in the previous summary, six drift tests were conducted with the ESG to verify the mathematical model calibration procedure developed from the simulation study. Additional drift tests were conducted to evaluate the mathematical model torque coefficients' sensitivity to changes in the suspension preload voltages, the rotor speed, and the gravity vector. Two of the six original tests were used as control tests for the second test phase. These two tests gave the best fit of the mathematical model to the data for the calibration procedure. These two tests had identical test conditions except for the spin axis location. The spin axis locus was 60 deg from the earth's polar axis (EPA) for the first case, and for the second case was 120 deg from the EPA. The Y force axis was northeast; X force axis was southeast; and the Z axis was vertical.

Six tests were conducted to evaluate the gravity vector influence on the ESG. Three groups of two tests were made, each group repeating the spin axis loci used in the control test. The first group of tests was conducted with the Z axis up, the second group was made with the Y axis up, and the third set with the Y axis down. A group of two tests was conducted to evaluate the effect of changing the suspension voltage. The two control tests were again repeated with the suspension voltage at the 15-g rather than the 4-g preload level of the control tests. One of the two tests was conducted with the spin-axis locus in the plane of the housing split and parallel to the earth equatorial plane. This test was then used as a prediction test for the gravity vector mathematical model solution for the drift coefficients.

Two additional tests were conducted repeating the control tests but with the rotor speed reduced from 200 to 180 rev/s. These were the only tests made at reduced rotor speed. All other tests were conducted at the 200-rev/s rotor speed level.

According to theory, the mass unbalance and third-harmonic torque coefficients are linear functions of the gravity vector, while the fourth harmonic and ellipsoidal rotor torque coefficients are quadratic functions of the gravity vector. Comparison of the torque coefficients

from the tests where the Y axis was up, with the torque coefficients from the tests where the Y axis was down, showed an indication of agreement between ESG torque theory and experimental results. A set of torque coefficients developed from these four tests was used for a prediction of the drift error compensation. The ratio of residual error to uncompensated error was approximately 25%, based on this prediction.

The results of the testing showed no preload sensitivity of the mass unbalance and the third harmonic rotor shape torque coefficients. The sensitivity of the fourth-harmonic rotor shape torque coefficient to the force voltage was determined to be proportional to the square of the force voltage. These results are in agreement with the ESG torque theory. However, the rotor translation torque coefficents were found to increase with increased preload, which does not agree with theory. This noncorrelation was attributed to electrode assembly errors.

The change in rotor operating speed had no discernible effect on the mass unbalance or on the third-harmonic rotor shape torque coefficients. A significant speed sensitivity was noted in the changes to the ellipsoidal rotor shape coefficients at the reduced rotor operating speed. These changes to the torque coefficients are in agreement with ESG torque theory.

F. Obstacle Detector for Surface-Landed Vehicles, *R. H. Burns*

1. Introduction

Work has begun on the development and evaluation of an automatic optical obstacle-detector for a lunar or planetary roving vehicle. Such a vehicle, particularly one for planetary use, should be fairly autonomous because of the long time lag involved in control from earth. Successful navigation will require a sensor that can detect possible obstacles such as rocks, holes, and impassable slopes, and relay their type and location to an on-board computer.

The sensor under development will perform this function by scanning the area in front of the vehicle with a laser radar that will yield range data as a function of azimuth and elevation. Rocks and holes will be identifiable by the range variations and discontinuities observed as the beam sweeps over them.

2. Design Objectives

Preliminary performance goals are as follows:

(1) A maximum range at least 100 ft (ultimately 500 ft) to allow sufficient warning to maneuver easily around larger obstacles.

(2) A scanned field of view covering at least the width of the vehicle out to the maximum range.

(3) A range resolution of less than 1 ft, in order to resolve the smallest obstacles that could affect the operation of a small roving vehicle.

(4) An angular resolution equivalent to about 1 ft at 50 ft, for the same reason.

(5) A capability of scanning the above field with the above resolution at least once a second, since the data will be used on a vehicle moving about 1 ft/s.

(6) Immunity from the motions of a vehicle traveling on a relatively rough terrain.

(7) Adherance to the standard flight hardware objectives of low power drain, low size and weight, and high reliability.

3. Optical Radar Transmitter

The development of an efficient source of light pulses less than 1 ns duration with a repetition rate of approximately 1 Mhz is in progress. These parameters are derived from the range resolution and maximum range goals. The source under study is a stored-pulse variation of a mode-locked helium–neon laser.

In a mode-locked laser a time-varying loss or phase shift is introduced into the optical cavity. When the loss or phase variation occurs at a rate synchronous with the round-trip time in the cavity, the optical output is bunched into pulses, which for a helium–neon laser are typically of 0.5 ns duration. The bunching does not affect the average power. Many devices have been used to cause mode-locking, such as piezoelectrically driven mirrors and saturable dyes. One device, attractive because of its potential simplicity and longevity, is a tube of excited neon gas.

A simple mode-locked laser yields pulses that are unsuitable for ranging because they are so close together in time. A laser 2 ft long will emit pulses about 4 ns apart. If a string of such pulses were used for ranging, the range data from an object 10 ft away would be identical to that

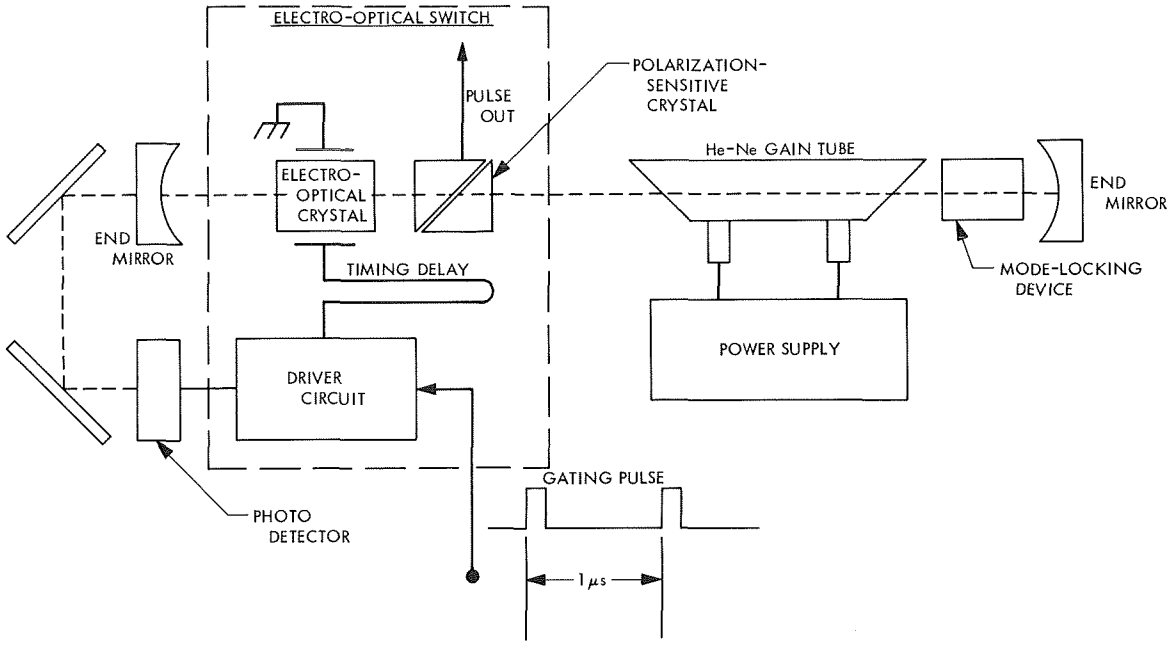


Fig. 11. Pulse-storage laser schematic

from an object 12, 14, or 16 ft away. The close spacing of pulses would cause an ambiguity. In order to get useful ranging data at distances up to 500 ft, it would be desirable to store the output of the laser for almost a microsecond and then emit it all at once. This technique was first demonstrated by Zitter, et al. (Ref. 1).

The essential parts of a pulse-storage laser are shown in Fig. 11. The unit consists of a very low-loss optical cavity containing a helium–neon gain tube, a mode-locking device, and an optical switch consisting of an electro-optical crystal, a polarization-sensitive prism, and a driver circuit for the crystal.

At the beginning of a pulse period, the unactivated optical switch has no effect and the laser begins to oscillate. The mode-locking device causes the resulting light to bunch into one pulse 4 to 6 in. long that reflects back and forth between the mirrors at the ends of the optical cavity, gaining amplitude with each pass through the gain tube. Little energy is lost at the mirrors because they have extremely high reflectance. A small amount does pass through and, when coincident with an electrical gating pulse, is used to trigger an electrical pulse applied to the electro-optical crystal. The gating pulse occurs at the end of each build-up period. The pulse applied to the crystal reaches it shortly after the optical pulse has passed through on its way to the gain tube. Full voltage is applied by the time the light returns. The light passes through the crystal, bounces off the end mirror, passes again through the crystal, and is thus converted to the opposite polarization. Striking the prism, it is deflected from the cavity in the form of a sub-nanosecond pulse containing most of the optical energy generated by the gain tube in the pulse build-up period. The build-up period is typically 1 μs; thus the peak pulse power is on the order of 1000 times the continuous wave power of the laser.

4. Experimental Results

The emphasis of the laboratory work so far has been on the establishment of a laser laboratory facility with the necessary high-speed instrumentation and on study and selection of the mode-locking device. The gain tube is provided by a Spectraphysics Model 124 helium–neon laser. It was initially expected that the nonlinear device would be a discharge tube filled with low-pressure neon gas (Ref. 2). Several tubes were designed, built, and tested. With certain cavity configurations, however, the laser has been found to be self-locking. Apparently under these conditions, now being studied, the gas in the helium–neon gain tube provides the nonlinearity necessary for a stable mode lock.

Figure 12 illustrates the mode spectrum and the pulse train achieved in a stable self-mode-locking configuration. Detector response time is thought to broaden the displayed pulse by approximately 1 ns. Peak pulse power is approximately 1/10 W.

References

1. Zitter, R. H., Steier, W. H., and Rosenberg, R., "Fast Pulse Dumping and Power Buildup in a Mode-Locked He–Ne Laser," *IEEE J. Quant. Electron.*, Vol. QE-3, pp. 614–617, November 1967.
2. Fox, A. G., Schwarz, S. E., and Smith, P. W., "Use of Neon as a Nonlinear Absorber for Mode Locking a He–Ne Laser," *Appl. Phys. Let.*, Vol. 12, pp. 371–373, June 1968.

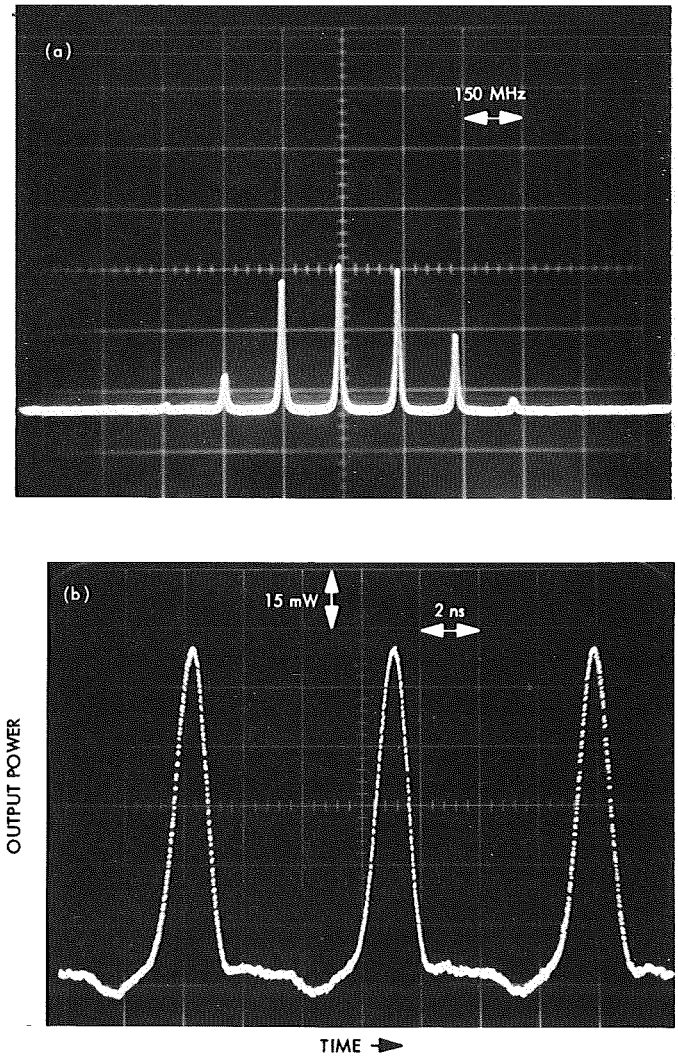


Fig. 12. Self mode-locked laser operation: (a) Mode spectrum, (b) Output power versus time

G. Computer-Animated Spacecraft Dynamics,

G. E. Fleischer

1. Introduction

While techniques for calculating the detailed motion of continuous dynamic systems have long proved themselves sophisticated, effective, and therefore valuable, to the modern designer/analyst, the capabilities for visually displaying the results of such computations often receive relatively little consideration. Short-term economics probably plays a greater role than any other factor in the utilizing of computer time to produce numerical data rather than to produce sophisticated visual aids. However, the objectives of computer simulation include not only the production of numerical answers but the creation of an insight and understanding, on the part of the analyst, of the system under study. Whereas the latter objective is of primary concern, a visual display of simulation results proves invaluable.

This report summarizes some recent efforts to produce an animated 16-mm silent film depicting the motion of the *Mariner IV* spacecraft in the days following depletion of its attitude-control gas. In particular, the intent was to show the result of attempts to realign the craft's roll axis toward the Sun by commanding roll gyros on and off periodically.

2. Animation System

For some time a Stromberg-Carlson SC-4020 microfilm recorder has been used at JPL in conjunction with IBM 7094 and Univac 1108 computers to provide graphical output either in the form of 35-mm microfilm frames or photographically enlarged hard copy (F-80). The recorder consists basically of a cathode-ray display tube and a camera. It is capable of executing simple instructions, such as advancing the film, displaying characters or spots at specified coordinates on the display tube surface, or drawing short line segments from one spot to another. The plotting surface is a 1024×1024-element grid and the device can draw points, lines, or characters at rates up to 100,000/s. Its potential for displaying continuous motion quickly and efficiently seems to rely chiefly on the existence of sufficiently powerful programming techniques.

The procedure used in this particular case is as follows: Instructions for creating the images to be plotted, frame-by-frame, in sufficient number to allow the visual impression of continuous movement, are submitted to the computer (Univac 1108) as a deck of punched cards. The language used is Fortran V. In addition, it is necessary to make use of a special library of plotting subroutines to activate the SC-4020. Execution of the Fortran program results, then, in the creation of a magnetic tape recording containing all the information necessary to produce the desired images when the tape is read by a

SC-4020 film recorder unit. After exposure and chemical developing, the 35-mm film strip may be shown directly using 35-mm projection equipment; or, since 16-mm equipment is more likely available, it may undergo further photographic processing and reduction to 16-mm stock.

3. Coordinate Systems and the Spacecraft Representation

Three distinct coordinate frames are used in the program to generate an animated sequence of *Mariner IV*'s motions. First, a spacecraft body-fixed frame is defined to allow simple descriptions of the vehicle's shape (assumed rigid) in terms of connected line segments. Second, an inertial frame is defined in which certain celestial references are assumed to be fixed, i.e., the Sun, Canopus, etc. It is this frame in which spacecraft motion will be described. Finally, a coordinate frame fixed to the observer is specified, providing the opportunity to view the motion from any vantage point. Transformation from coordinates specified in the body-fixed frame to those of the inertial frame, or vice-versa, is provided by the familiar 3–1–3 Euler angle sequence. Angles ϕ, θ, and Ψ, respectively, represent the sequence of rotations about body axes 3, 1, and 3 (z, x, and z) necessary to uniquely arrive at the body's orientation with respect to the inertial reference axes. Thus the inertial coordinates x_I, y_I, and z_I of a point on the craft are given in terms of body coordinates x_B, y_B, z_B from

$$\begin{Bmatrix} x_I \\ y_I \\ z_I \end{Bmatrix} = [B] \begin{Bmatrix} x_0 \\ y_0 \\ z_0 \end{Bmatrix}$$

where

$$B_{11} = \cos\Psi\cos\phi - \sin\Psi\sin\phi\cos\theta$$

$$B_{12} = -\sin\Psi\cos\phi - \sin\phi\cos\Psi\cos\theta$$

$$B_{13} = \sin\theta\sin\phi$$

$$B_{21} = \sin\phi\cos\Psi + \sin\Psi\cos\theta\cos\phi$$

$$B_{22} = -\sin\phi\sin\Psi + \cos\Psi\cos\theta\cos\phi$$

$$B_{23} = -\sin\theta\cos\phi$$

$$B_{31} = \sin\Psi\sin\theta$$

$$B_{32} = \cos\Psi\sin\theta$$

$$B_{33} = \cos\theta$$

A similar sequence of rotations ϕ_P, θ_P, Ψ_P may be defined to locate the observer's frame with respect to inertial coordinates:

$$\begin{Bmatrix} x_I \\ y_I \\ z_I \end{Bmatrix} = [A] \begin{Bmatrix} x_0 \\ y_0 \\ z_0 \end{Bmatrix}$$

where the elements of A are identical to those of B after substitution of ϕ_P for ϕ, θ_P for θ, and Ψ_P for Ψ.

Figure 13 indicates the position of the observer in the observer's frame of reference. Since the origins of all three reference frames are assumed coincident and the figure to be centered at the origin, the observer is placed on the $-z_0$ axis an arbitrary distance d_0 from the origin. The picture plane, upon which the desired two-dimensional perspective view of the body will be projected, is located a distance d_P from the origin along $-z_0$.

Projection of the craft's image on the picture plane is accomplished by the following steps:

(1) Coordinates defined in the body-fixed frame are transformed to the inertial frame and then to the observation frame, i.e.,

$$\begin{Bmatrix} x_0 \\ y_0 \\ z_0 \end{Bmatrix} = A^{-1}B \begin{Bmatrix} x_B \\ y_B \\ z_B \end{Bmatrix}$$

(2) Any coordinates defined in the inertial frame are transformed to the observation frame.

(3) Finally, image coordinates in the picture plane, including the effects of perspective, are obtained from

$$x_P = x_0 \left(\frac{d_P}{d_0 + z_0} \right) K$$

$$y_P = y_0 \left(\frac{d_P}{d_0 + z_0} \right) K$$

where K is an arbitrary enlargement factor which may be used to adjust the image size, if necessary.

An image, in perspective, of a *Mariner*-type spacecraft resulting from the computational steps outlined above is given in Fig. 14. All connected line segments used to describe the body are, of course, visible. Included is an arrow, defined in inertial coordinates, as pointing toward the Sun. To achieve the effect of motion, successive frames

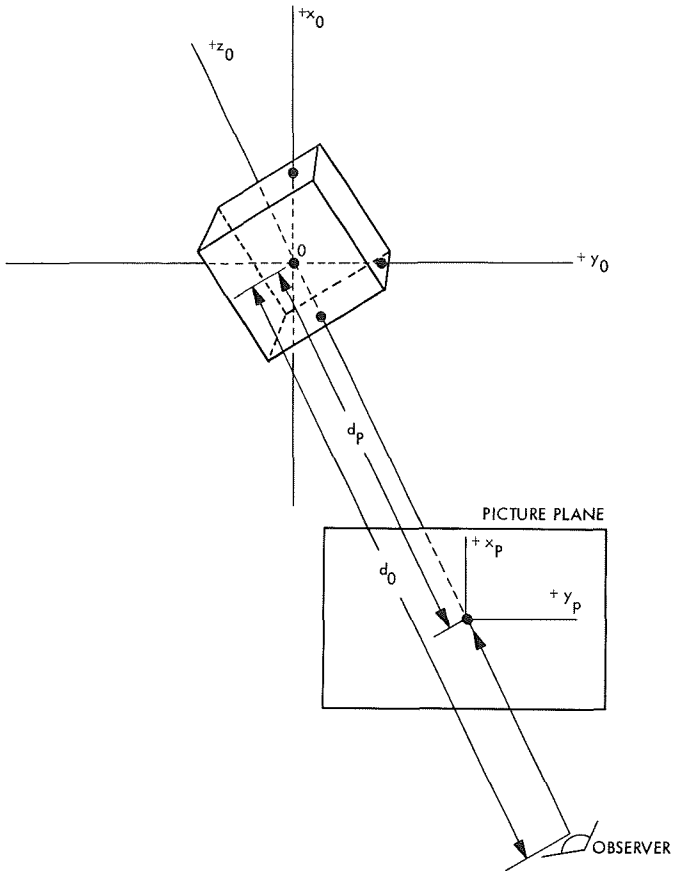


Fig. 13. Observer's coordinate frame x_0, y_0, z_0

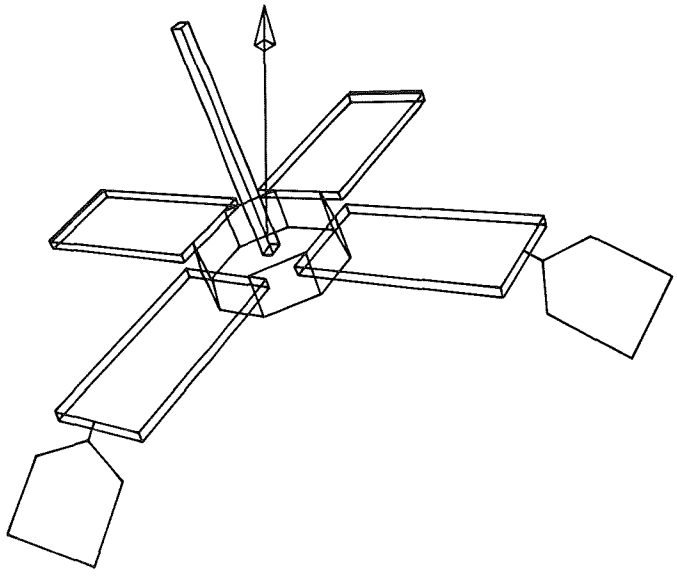


Fig. 14. Computer-generated image of *Mariner IV* with all edges visible

of the spacecraft image in the picture plane will correspond to an input sequence of ϕ, θ, and Ψ values describing body motion in the inertial frame. The Sun-directed arrow remains fixed in each view, providing some indication of the frame of reference.

4. Hidden-Line Removal

It is possible, regardless of the helpful effects of perspective plotting, to view the image in Fig. 14 in two different ways. The correct view being that the omni-antenna emanates from the paper, the other impression being that the antenna enters the paper, in spite of the foreshortened panels. And, if these visualizations suddenly happen to switch while one is viewing rotation of the craft, the illusion of a sudden reversal of rotation is created. Further, it is often difficult to force one's mind to reverse the visual illusion effect once it has been adopted.

To avoid any possible confusion, it is clearly desirable to eliminate from view those edges of the craft which will be hidden from the observer if the body is to be considered a three-dimensional solid. The approach taken to achieve this begins by requiring that all line segments defining one, two, or three-dimensional objects in either body-fixed or inertial coordinates be uniquely labeled. In addition, each plane formed by groups of line segments must be uniquely labeled and defined by supplying, as part of the program's input data, the set of labels of those line segments which form each plane. Thus, the sequence of data cards defining planes might contain the following information:

Card 1: 052

Card 2: 5, 6, 33, 56

Card 3: 9, 11, 26, 28, 29

etc.

The first card indicates the number of planes to be considered by the program. The second card defines plane 1 as being made up of line segments 5, 6, 33, and 56. Card 3 similarly defines plane 2, etc. It is also convenient to order the line segment labels such that if one traveled around the plane's borders, from line to line, in that order, the direction of rotation would be that of a right-handed screw normal to the plane but *emerging* from that side of the plane which is an *outer* surface of the solid body. In this way, the program can quickly determine by examination of such a normal that the plane's outer face cannot be seen by the viewer, and therefore need not be considered as a possible obstruction to viewing parts of the vehicle. The four solar vanes

of Fig. 14 don't have "inner" and "outer" sides and therefore must always be included as potential obstructing planes. Special instructions must be introduced to take care of such situations.

The procedure to eliminate hidden lines generally takes the following form: A discrete number of points along each line segment, from end to end, are tested to determine whether or not they are hidden. The algorithm used to test each point sifts through each of the planes that could possibly obstruct the observer's view of it. Each of these planes has previously been examined to determine its maximum and minimum coordinates, x_P and y_P, in the picture plane. If the test point's picture plane coordinates fall outside these boundaries, the plane cannot hide it, and a new plane is examined. Further tests will be made if the point is within the boundaries, such as:

(1) If the point's z_0 coordinate is less than plane's smallest z_0 coordinate, it can't be hidden by this plane.

(2) If the point is on a line which forms the boundary of the plane, it can't be hidden by this plane.

Certain characteristics and symmetries of a specific solid figure will enable other specialized tests to be made as well. However, after all previous test conditions have failed to eliminate a plane as an obstruction, a final definitive test is made which determines the location of the intersection of the observer-to-point line-of-sight and the plane in question. If this intersection lies within the boundaries of the plane and the point is beyond the intersection, the point must be hidden. Figure 15 pictures the result of applying the hidden-line algorithm to another view of *Mariner IV*.

It is also easy to add or delete lines to the generated picture as a function of certain input data and thereby achieve additional illustrative effects. For example, as mentioned earlier, the animated motion of *Mariner IV* was intended to show the effects of turning roll gyros off and on. Figure 16 shows the addition of an arrow used to signal the application of momentum due to gyro turn-on. A few appropriate program statements allowed the arrow to be plotted or deleted, depending on the appearance of an indicator digit on an input data card.

5. Future Applications

The production of a 16-mm silent film incorporating approximately 3000 computer-generated frames has been

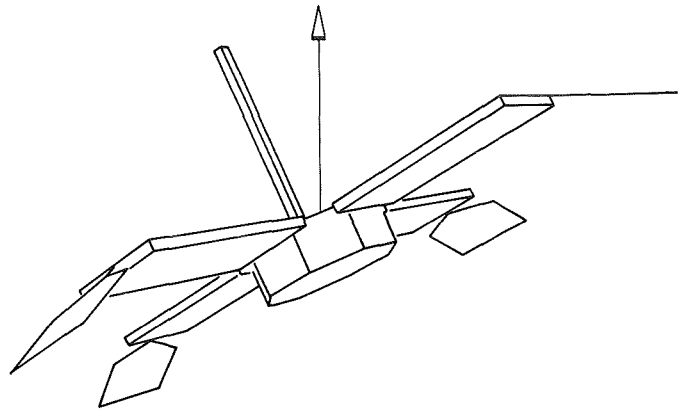


Fig. 15. Computer-generated image of *Mariner IV* with hidden edges removed

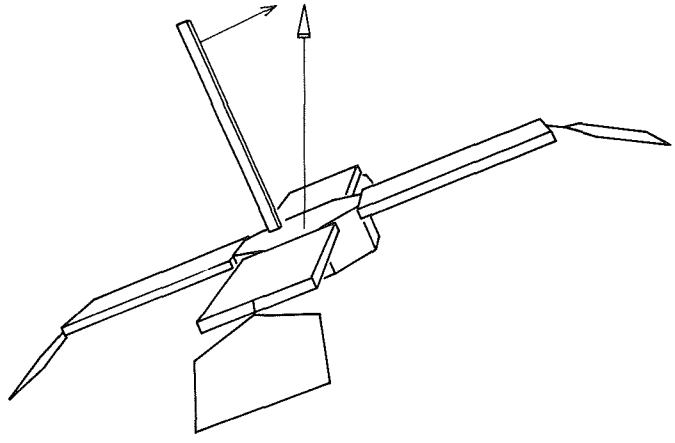


Fig. 16. Application of momentum to *Mariner IV* illustrated by conditioned plotting of a two-dimensional arrow

successfully completed. The *Mariner IV* spacecraft representation included more than 120 lines and 49 planes. From 8 to 20 points along each line were examined by the hidden-line algorithm. Univac 1108 processor time per frame varied between 2 and 4 s, depending on the particular view required. The hidden-line algorithm can almost certainly be improved to reduce this time to less than 1 s. For instance, although this was not done for lack of time, rather than make a sometimes inconclusive test to determine if the test point's picture-plane coordinates are within or beyond a plane's *maximum* and *minimum* picture-plane coordinates, it is possible, using sets of inequalities, to determine absolutely whether it is in or out of the plane's boundaries in the picture plane.

It is impossible to foresee just how such a pictorial tool might be used in various circumstances. However,

it could prove valuable as a real-time display for SFOF personnel. Looking ahead to the time of multi-terminal 1108 graphics displays, the film medium might be abandoned completely in favor of semicontinuous TV pictures to aid systems designers and analysts. This, in fact, has already occurred in several other research installations across the country.

H. A Digital Momentum Wheel Control System for Spacecraft Attitude Control, *W. E. Crawford*

1. Introduction

The requirement for highly accurate antenna pointing on future multi-planet missions has placed severe requirements on the attitude-control system. Not only must the spacecraft attitude be controlled one order of magnitude more accurately than that of the *Mariner* Mars 1969, *Mariner* Mars 1971, and *Viking* spacecraft, but the attitude control must be able to bias the spacecraft Z axis away from the Sun-line toward the earth and still maintain pointing accuracy. A wide-angle digital Sun sensor with an incremental resolution of 0.025 deg has been developed and is now being tested. Electronic circuits to process the digital Sun sensor position information and also process the position bias are now being breadboarded.

Figure 17 is a block diagram of the digital signal summing circuit showing the Sun sensor, central computer bias, and gyro input signals along with the digital processor necessary to sum the signals.

Figure 18 is a block diagram of the digital momentum wheel control electronics. It is the mechanization of the circuitry in this second diagram that will be discussed in this report.

2. Circuit Function and Design

The usual method for actuator (momentum wheel) control in a digital system is to convert the digital signal at the output register to an analog signal (digital-to-analog [D/AC] converter), and send this signal through a power amplifier to the actuator. This technique is not a good solution for the present system because of the huge quantization error in the digital-to-analog conversion. The total allowable position deadband is only

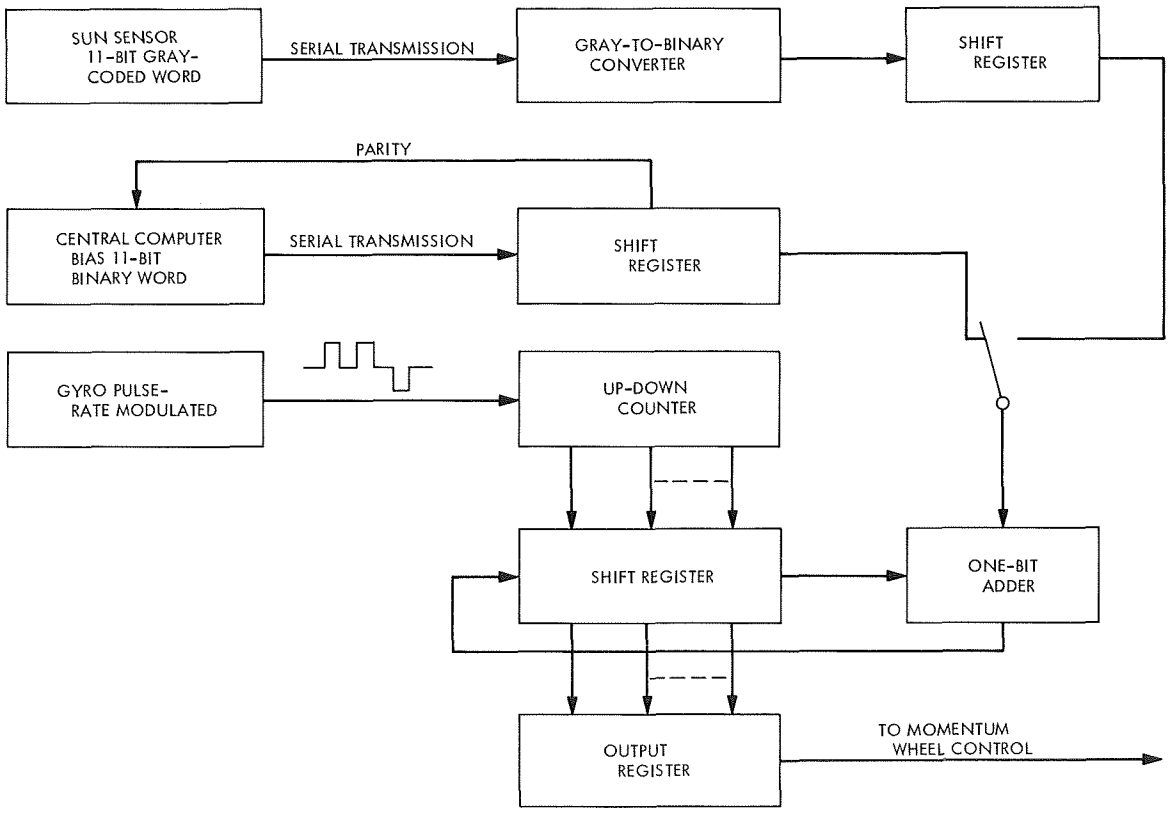


Fig. 17. Digital signal summing circuit

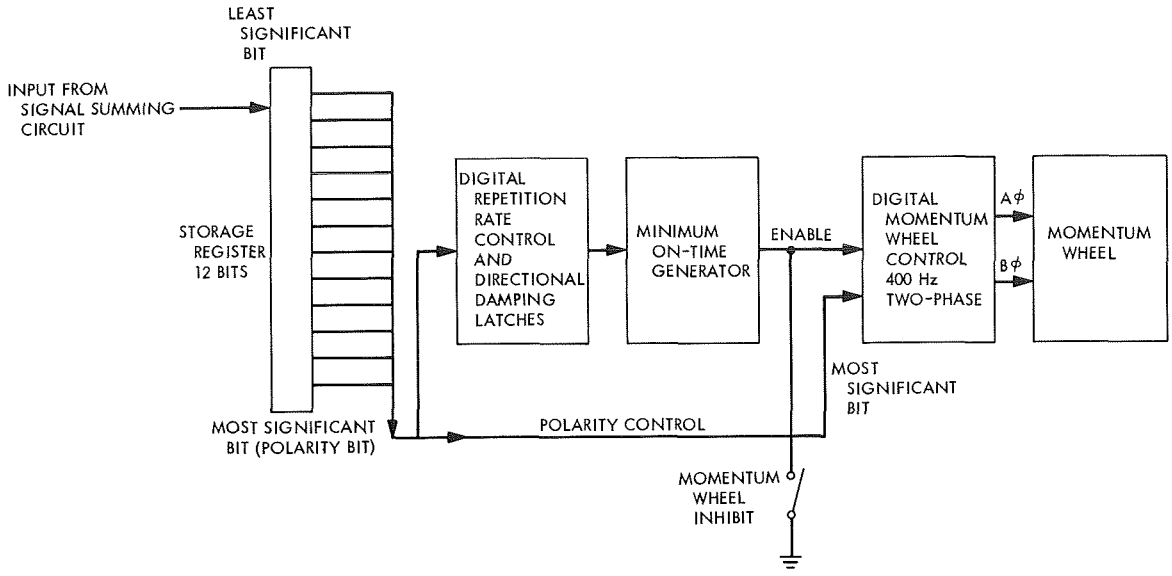


Fig. 18. Momentum wheel control system block diagram

0.05 or $\pm$0.025 deg. This is exactly two increments of Sun sensor position for the complete deadband. This means that the system is inherently digital whether or not it is converted to analog. Since 11-bit D/ACs have large dc errors when compared to the least significant bit, it is necessary to control the momentum wheel using digital techniques.

A method for digital control is shown in Fig. 18. During the cruise mode of operation, when the gyros are not on, a method for creating rate information must be generated. This is analogous to derived-rate attitude control in the *Mariner* spacecraft. In order to accomplish derived rate in a digital system, the mechanization shown in Fig. 2 was instrumented.

Stated briefly, the design philosophy of the digital momentum wheel control electronics is as follows:

(1) The momentum wheel is to be energized at full voltage for a fixed length of time (a minimum-on-time or bang-bang system).

(2) As the storage register error (position error) gets larger, the digital repetition pulse rate increases. Each pulse causes a momentum wheel actuation.

(3) As the storage register error gets smaller, the digital repetition pulse rate decreases.

(4) As the storage register error gets larger, each bit change causes a minimum-on-time momentum wheel actuation by energizing the directional damping latches.

(5) As the storage register error decreases, no directional damping latches are set.

(6) As the digital error changes polarity (most significant bit change), the directional damping latches are reset.

(7) Each pulse at the minimum-on-time generator input causes the generator to turn *on* and/or remain *on* for a minimum-on-time interval subsequent to the pulse.

The control system consequently behaves in the following manner: The restoring torque supplied by the momentum wheel to the spacecraft is greater when the spacecraft is moving away from null and less when the spacecraft is moving toward null. This is the principle of directional damping and will cause the spacecraft to converge to the null point of the biased Sun sensor.

3. System Performance

Figures 19, 20, and 21 show how the system will behave under different cruise mode conditions. Figure 19 shows spacecraft position and rate in the presence of a large momentary disturbance (meteorite impact, etc). The ability of the system to return to null is clearly shown. Figures 20 and 21 show how the control system will behave in the limit cycle (within the deadband). Figure 20 shows the ideal behavior of the spacecraft with no solar torque or external disturbance torque. Figure 21 shows the system behavior in the presence of

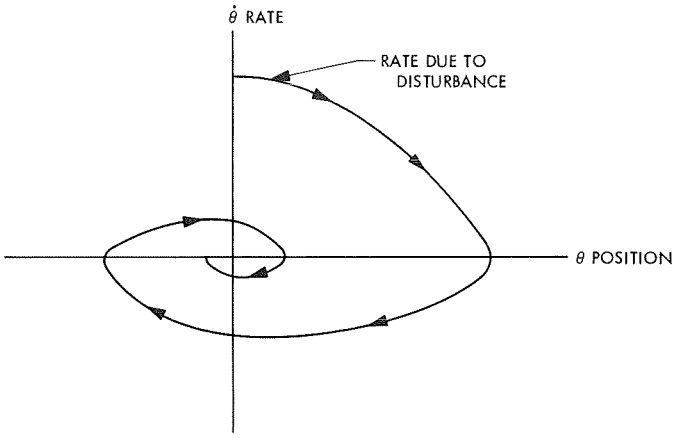


Fig. 19. Effect of momentary disturbance torque on digital momentum wheel control system

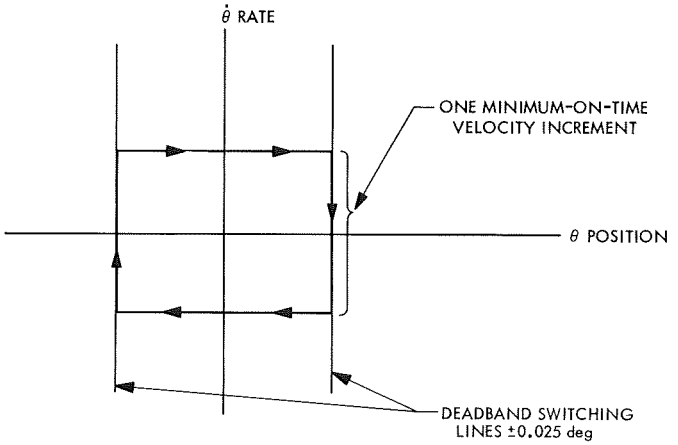


Fig. 20. Ideal limit cycle for digital momentum wheel control system

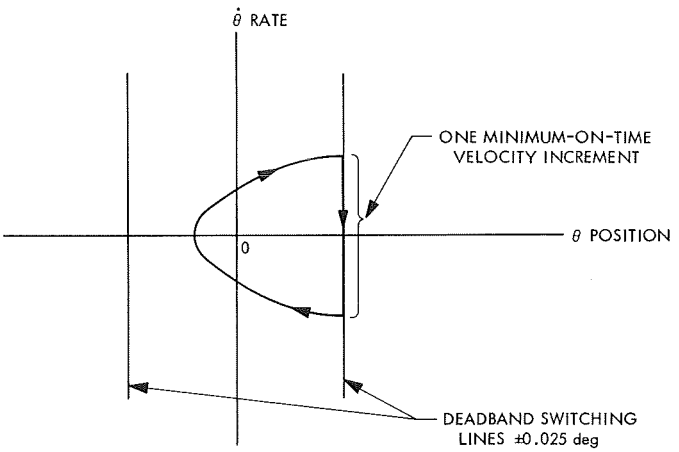


Fig. 21. Effect of continuous disturbance torque on digital momentum wheel control system limit cycle

a small disturbance (solar torque or outgassing components).

4. Conclusion

The electronic circuitry described in this report is now being breadboarded and tested for use in the *Thermoelectric Outer Planet Spacecraft* single-axis air-bearing table demonstrations. Mating of the digital electronics with the momentum wheel will be performed in early 1970. Digital circuit components are almost all exclusively Texas Instruments low-power 54/74 L integrated circuits. Since the momentum wheel electronics circuitry has a large number of discrete components, there is a good possibility of incorporating hybrid microcircuits in the design in order to decrease volume and increase reliability.

XIII. Guidance and Control Research

GUIDANCE AND CONTROL DIVISION

A. Ion Sputtering of Oxygen-Contaminated Tantalum, *S. S. Luebbers*

1. Introduction

As an extension of the work reported in SPS 37-55, Vol. III, pp. 132–136 and Ref. 1, the electron emission properties of an oxygen-contaminated tantalum wire probe have been studied for the surface conditions obtained after various degrees of cesium-ion sputtering. The experiments were performed in a plasma-immersion-probe tube where ion sputtering was accomplished *in situ* by applying a negative 50-V bias to the experimental probe. With the probe thus biased, cesium ions formed in the main discharge are accelerated toward the negative probe and sputter the surface of adsorbed gas atoms. By studying the electron emission from the cesiated electrode, we are able to qualitatively determine how the surface is affected by ion sputtering. We find that the low-voltage (50 V) cesium-ion sputtering is effective in removing surface oxygen to a certain characteristic depth. After several hours of probe operation at 1000°K, the electron-emission current from the probe increases to a value near that observed prior to sputtering. Experiments have been performed to establish the oxygen-diffusion behavior within the tantalum-probe material.

2. Experimental Techniques

The plasma-anode technique for studying the emission properties of cesiated metals has been discussed extensively in SPS 37-55, Vol. III and Refs. 1–3. A complete description of the experimental tube and test circuitry used in these experiments may be found in Ref. 1. The experimental tube (plasma-immersion-probe tube) consists of an electrically heated tungsten filament and a large anode disk located at opposite ends of a cylindrical pyrex glass envelope 3 in. in diameter and 12 in. long. The tube is filled with liquid cesium which, when heated in a controlled oven, establishes an equilibrium vapor pressure of cesium in the tube. The cesium vapor is ionized to produce a glow discharge by an applied potential between the filament and anode. The test probes, whose emission properties we wish to determine, are inserted into the discharge from a side port, and volt–ampere curves are measured for different probe and cesium temperatures. The cesium vapor and accompanying discharge modify the probe work function and neutralize space-charge effects that

result from electron emission from the hot test probe. The probe is sized to ensure that the emission current is only a small perturbation on the total discharge current of the tube. The probe is electrically heated by half-wave pulsating current with the voltage–current measurements made during the off half cycle. The thermionic electrons from the test probe are emitted into the plasma column of the discharge that acts as the collecting anode.

From the measured volt–ampere curves, the electron-saturated emission from the probe is established and related to the surface work function via the Richardson equation. In this manner, we are able to study the test-probe work function behavior and its relationship to ion sputtering.

3. Test-Probe Construction

Details of the test-probe structure are shown in Fig. 1. The test-probe material is 10-mil tantalum wire containing bulk-absorbed oxygen. The oxygen content was determined by neutron-activation experiments to be approximately 0.12 at. %. Alumina insulator tubes define the probe emission area and support the probe wire. The insulator extends into the tube side port where it is attached to the glass press joined to the main tube envelope. A "platinum bright" guard ring (not shown) is fired onto the tube glass wall at the base of the test-probe structure to collect leakage currents originating from the tube walls. As indicated in the figure, the area of the exposed tantalum wire probe is 0.15 cm^2.

Prior to any emission measurements, a probe-temperature versus heating-current calibration was made over the visible range of a pyrometer and extrapolated into the low temperature region by using resistance measurements (Ref. 4). During actual emission measurements, probe temperature was determined by measuring the heating current since cesium discharge and filament radiation within the tube would enter into any pyrometric temperature measurement.

4. Work-Function Behavior

The work function of an uncontaminated metal in a cesium environment has been shown, both theoretically and experimentally, to be a unique function of the ratio of the metal surface temperature T_P to the liquid cesium-reservoir temperature T_{Cs} for a given metal having an uncesiated work function of ϕ_0 (Ref. 5). The theoretical cesiated work function for a material with $\phi_0 = 4.1$ eV is shown as a dashed line in Fig. 2. A plot of this type provides a convenient means to display the electron emission data from the experimental probe.

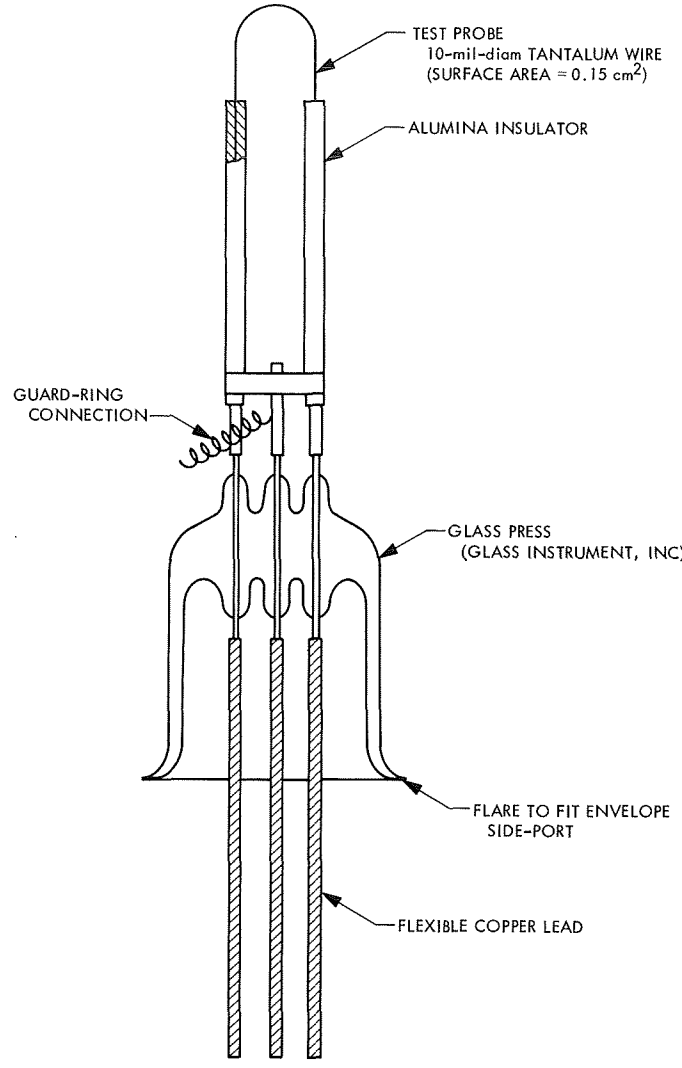


Fig. 1. Detail of test probe structure

5. Experimental Procedure

The plasma-immersion-probe tube was baked and evacuated to 10^{-8} torr and was subsequently charged with cesium. The results of initial emission measurements are shown in Fig. 2 indicated by the curve titled BEFORE FLASHING. At high temperature ratios, the data indicate that the bare emitter work function is approximately 4.4 eV. The experimental data have no correlation with the theoretical work function behavior of an uncontaminated material with a similar bare work function. This "non-ideal" behavior indicates that the surface structure is not simply a cesium-absorbed-on-tantalum structure.

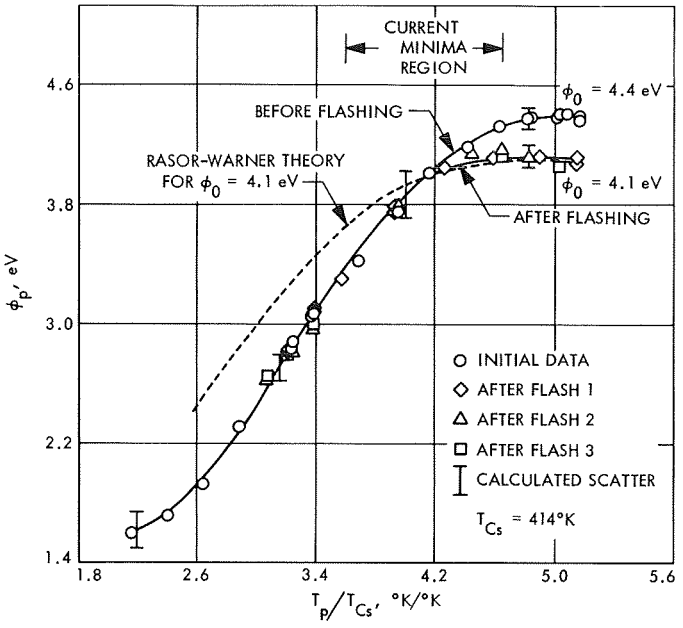


Fig. 2. Probe work function versus T_P/T_{Cs}

Repeated flashing of the probe, to temperatures as high as 2470°K, resulted in a shift of the low-cesium-coverage region (high T_P/T_{Cs}) of the curve to that predicted for a surface with a bare work function ϕ_0 near 4.1 eV, while in the higher cesium-coverage region (low T_P/T_{Cs}), the effect of flashing was negligible. This shift occurred immediately after the first flash, and further flashing caused no measurable change. The effect of the high-temperature flashing was to reduce the original concentration of oxygen within the bulk-tantalum probe to a value at which the rate of diffusion through the lattice was insufficient to replenish oxygen desorbed from the high-temperature surface; therefore, the anticipated work-function behavior for "clean" tantalum resulted in the high-ratio region. At lower temperature ratios, diffusion limitations were not observed and enhanced emission was still evident. Both the oxygen diffusion rate through the lattice and the evaporation rate from the surface are reduced by lowering the probe temperature; however, the reduction in evaporation rate is dominant since it depends exponentially upon temperature.

To see the importance of the work function deviation shown in Fig. 2, imagine a cesiated thermionic emitter operating at a temperature ratio $T_P/T_{Cs} = 3.0$ (a typical ratio for thermionic power converters). We note the experimental curves give a work function that is lower than theoretical value by 0.54 eV. At thermionic temperatures, we realize that an emission-current enhancement much greater than an order of magnitude results with this emitter material. We suggest that enhancement can be attributed to the fact that the test probe contains bulk-absorbed oxygen that acts as an emission-enhancing additive when it diffuses to the surface and coadsorbes with the cesium.

In Ref. 4, J. M. Houston and P. R. Dederick observed the same anomalous emission from a tantalum wire probe. Their experiments were performed in a similar tube but did not include the high probe-temperature region where we observed that flashing would reduce the work function to a consistent theoretical value. At the temperatures included in their measurements, agreement between the two different experiments is good. In their work, the effect of 105-V cesium-ion sputtering upon the electron-emission maximum was studied. The result of ion sputtering was to reduce the current maximum to a value typical of "clean" tantalum. After sputtering, the probe was repeatedly flashed at different temperatures and the current maximum was observed to gradually increase although the duration of flashing was not sufficiently long to allow the current maxima to increase to the originally observed value.

The test probe in our experiment was sputtered with 50-V cesium ions for 2 min at a probe temperature near 1000°K. After this initial sputtering, the work-function behavior shown by the solid line in Fig. 3 was obtained. To ensure a clean surface, a 30-s sputter was made immediately before each data point was taken. The results are typical of a clean tantalum surface with a bare work function near 4.1 eV and indicate the low-voltage sputtering was effective in removing the additive oxygen from the surface.

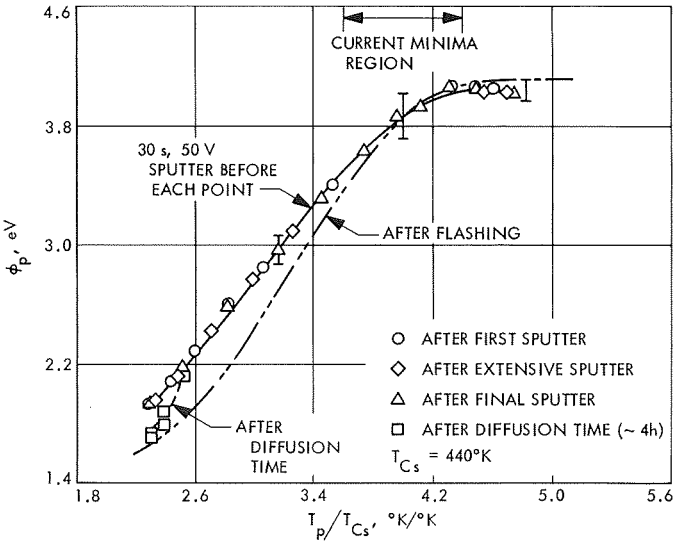


Fig. 3. Probe work function versus T_P/T_{Cs}

To verify that sputtering was only effective in removing oxygen from the bulk material in the immediate vicinity of the surface, diffusion experiments were performed. The test probe was left at a temperature of 1000°K and the current was recorded as a function of time. In Fig. 4, the complex, gradual increase in current with time is shown. This complexity may have resulted from the nonlinear buildup of oxygen on the surface with time, or a nonlinear behavior of work function at relatively high oxygen coverages, or by a combination of these effects. The simple exponential increase in current occurring for relatively short time duration (i.e., $\tau < t < 1400$ s) does, however, indicate that, initially, the work function of the emitter is being lowered linearly with time by the oxygen flux from within the bulk of the wire. (The initial slope of the curve is due to the incomplete removal of all oxygen within the depth effected by ion sputter.) In the low-oxygen-coverage region, where oxygen–oxygen interactions would not be present, we might expect the work function to decrease linearly with oxygen coverage. Considering this to be true, a diffusion constant D of L^2/τ is obtained where the average diffusion velocity is taken to be L/τ (L is the effective ion-sputtering depth). For effective ion-sputter depths of several atomic spacings ($L \sim 10^{-7}$ cm), an anomalously small diffusion constant ($D \sim 2 \times 10^{-17}$ cm^2 s^{-1}) is calculated. To obtain a reasonably large diffusion constant for oxygen in tantalum at 1000°K, the effective ion-penetration depth must be $>10^{-7}$ cm.

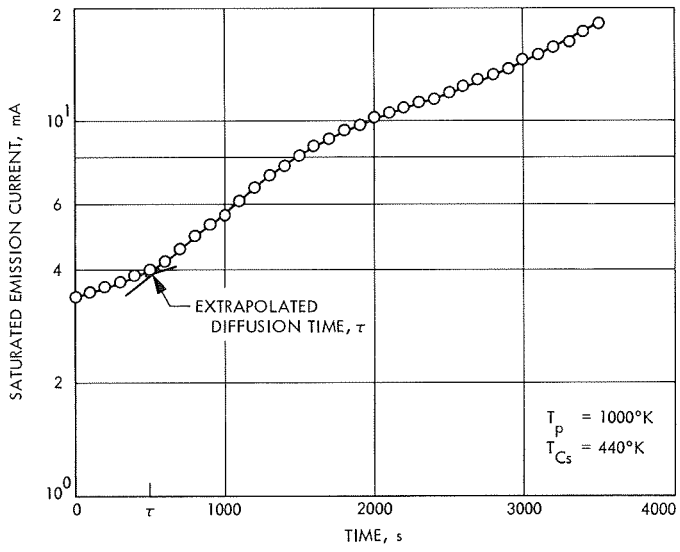


Fig. 4. Saturated emission current versus time (after sputtering)

Figure 5 shows a plausible oxygen-concentration profile at three different times after sputtering. The concentration profile in Fig. 5b would occur when the first group of oxygen atoms from within the bulk of the wire reaches the surface in time τ, while the last figure would be the equilibrium profile obtained after several hours. Under equilibrium conditions, the surface concentration of oxygen n_s, is maintained in the presence of desorption by the oxygen diffusion flux to the surface.

Similar experiments were repeated after different time lengths of sputtering. After each additional sputter, the slope of the curve in Fig. 4 for $\tau < t < 1400$ s was reduced

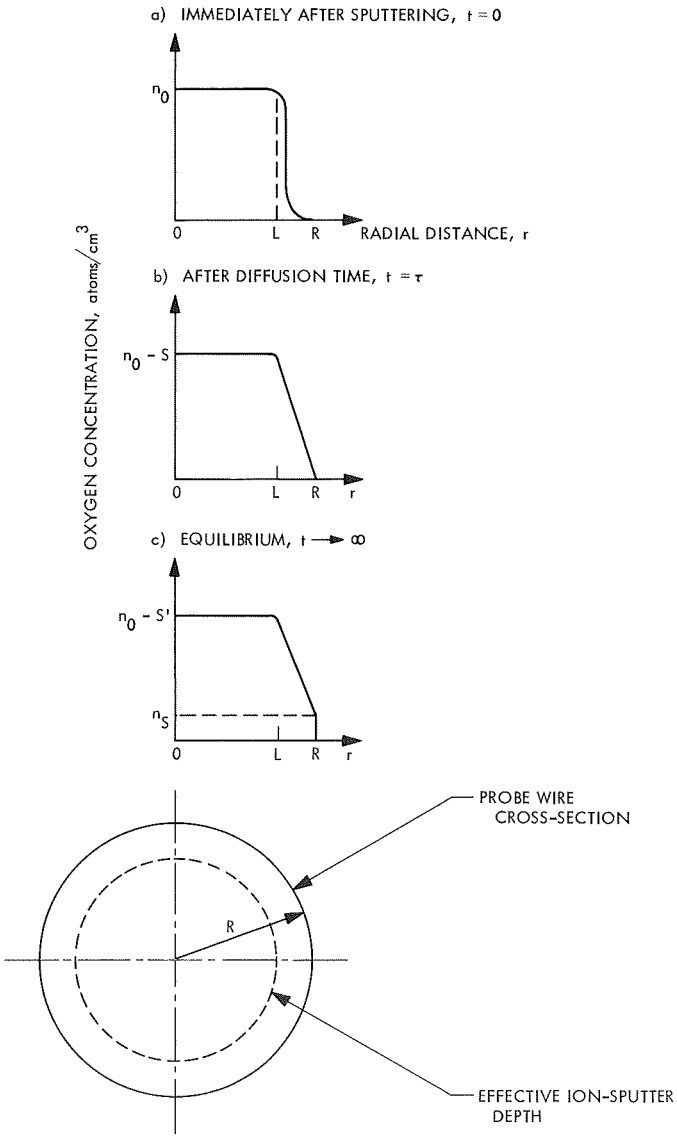


Fig. 5. Concentration profiles

significantly, indicating a decrease in oxygen driving force. There was also a tendency for the diffusion time τ to increase with time of sputtering. The current increase for $t > 1400$ s was not simply related to the history of probe sputtering; however, even after several hours of sputtering, the anomalously high emission current could be reestablished by maintaining the probe at 1000°K for an extended period of time (i.e., several hours).

One diffusion experiment was interrupted by an overnight shut-down of power at time $t = 1800$ s. After reheating the probe in the morning, the same emission current was observed and continued to increase until the test was terminated. Since the tube remained hot for an hour or so following shutdown, the return of the emission current to its value of the previous night indicates that the oxygen is not of tube-residual-gas origin.

After the current reached nearly a constant value, the probe temperature was increased and emission measurements were again made. Since we originally assumed that the sputtering was only effective in removing a small fraction of the oxygen (i.e., an effective ion-sputtering depth of several atomic layers), the work-function behavior was expected to return to that recorded after flashing only. However, this was not observed. Oxygen additive effects remained only at the low-temperature-ratio region where a small concentration of bulk oxygen would be sufficient to replace the oxygen evaporated from the surface. (Refer to the dashed line in Fig. 3 labeled AFTER DIFFUSION TIME.) This behavior further indicates that the low-voltage ion sputtering was effective in removing a significant quantity of oxygen from the wire probe and the effective ion-sputter depth was greater than several atomic spacings.

6. Conclusion

Although further experimental methods must be exploited to obtain a full and consistent picture, the behavior of a bulk-oxygen-contaminated emitter may be explained qualitatively. The diffusion of oxygen in tantalum provides an oxygen additive to an emitter surface which, when coadsorbed with cesium, gives enhanced emission. Low-voltage cesium-ion sputtering is effective in removing oxygen to a depth much greater than several atomic spacings.

References

1. Luebbers, S. S., and De Steese, J. G., "Oxygen Effects in Tantalum Thermionic Emitters," IEEE Report of the Thermionic Conversion Specialists Conference, p. 64, 1968.
2. Marchuk, P. M., *Trudy Inst. Fiz. Nauk. Ukraine*, Vol. 7, No. 17, 1956.
3. Houston, J. M., and Dederick, P. K., "Thermionic Emission of Thermionic Converter Collector Materials in Cs Vapor," IEEE Report on the Thermionic Conversion Specialists Conference, p. 77, 1964.
4. Houston, J. M., and Dederick, P. R., "Thermionic Emission of Various Metals in Cesium Vapor," Report 65-RL-3872E, General Electric Research Laboratory, p. 48–49, 1965.
5. Rasor, N. S., and Warner, C., "Correlation of Emission Processes for Adsorbed Alkali Films of Metal Surfaces," *J. Appl. Phys.*, Vol. 35, No. 9, pp. 2589–2600, Sept. 1964.

XIV. Applied Mechanics

ENGINEERING MECHANICS DIVISION

A. Multilayer Insulation Test Facility,

W. F. Carroll and L. D. Stimpson

The long-range objective of this work is to develop materials technology and fabrication processes necessary to optimize the performance characteristics of multilayer insulation systems for use on spacecraft. Such optimization includes weight and volume considerations and insulation efficiency, but principal emphasis will be on realistic fabrication and installation influences, and on shield predictability and reproducibility.

Techniques have been developed at JPL for the design and fabrication of practical thermal blankets for the *Ranger* and *Mariner* spacecraft. The degradation effects on thermal behavior of attachments, penetrations, seams, etc., dominate the shield performance but are not well understood. On the other hand, extensive experimental and analytical work has been performed elsewhere on ideal shields with minimized thermal conductance for cryogenic storage applications.

The calorimeter (Fig. 1) consists of a cylindrical test section with a guard section at each end. The guard sections are thermally isolated from the test section by low conductance spacers. The temperature of the guards is slaved to the test section by heaters controlled by differential thermocouples in the spacers. There are two etched circuit-type strip heaters in each section diametrically opposite each other and wired in series. The thermocouple and power leads are brought out through the lower guard to minimize conduction effects.

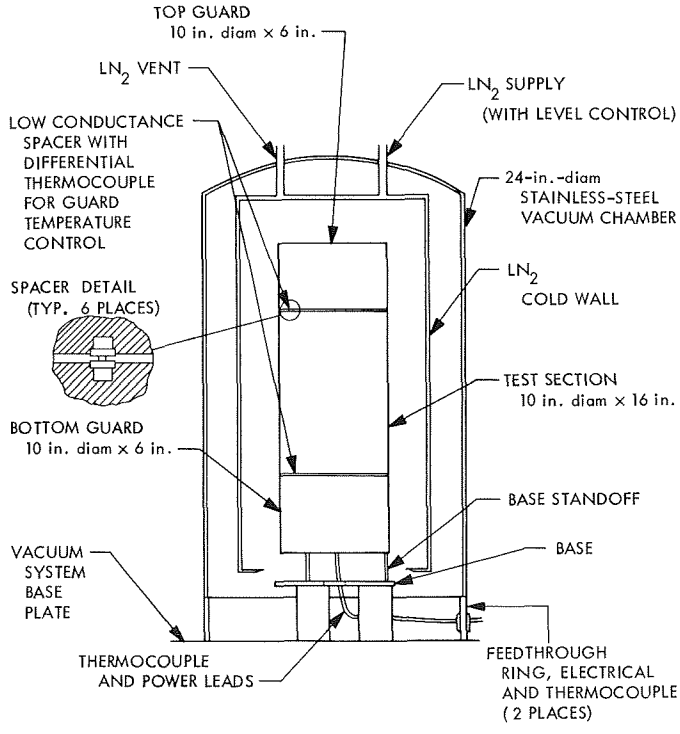


Fig. 1. Multilayer insulation thermal vacuum test setup

Figure 2 shows the insulation-wrapped calorimeter in the open vacuum chamber with the power and vacuum control panels below. Three separate precision power control units are visible on the lower left for controlling the electrical heating of the test and guard sections shown in Fig. 1. The lower lip of the liquid nitrogen (LN_2) cold wall can be seen inside the stainless-steel bell jar. The wrinkled pattern on the multilayer insulation is a characteristic that has been apparent after each test.

Figure 3 shows the existing instrumentation rack where both strip chart recorders and digital voltmeters are mounted. Most of this equipment was salvaged, but

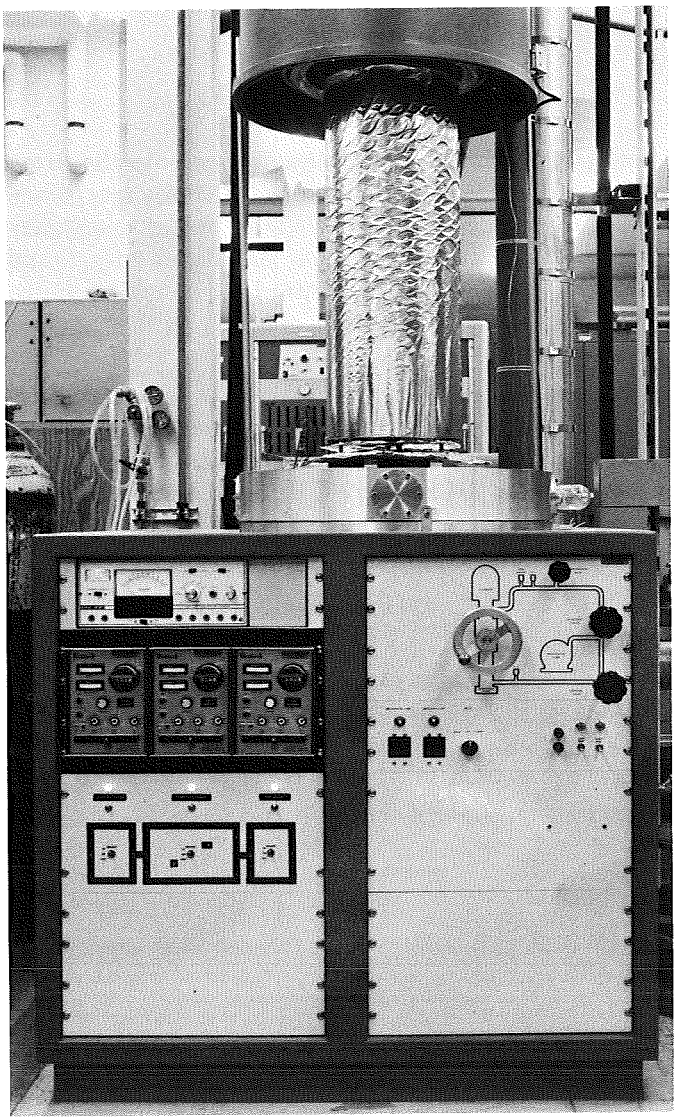


Fig. 2. Insulation-wrapped calorimeter in open vacuum chamber

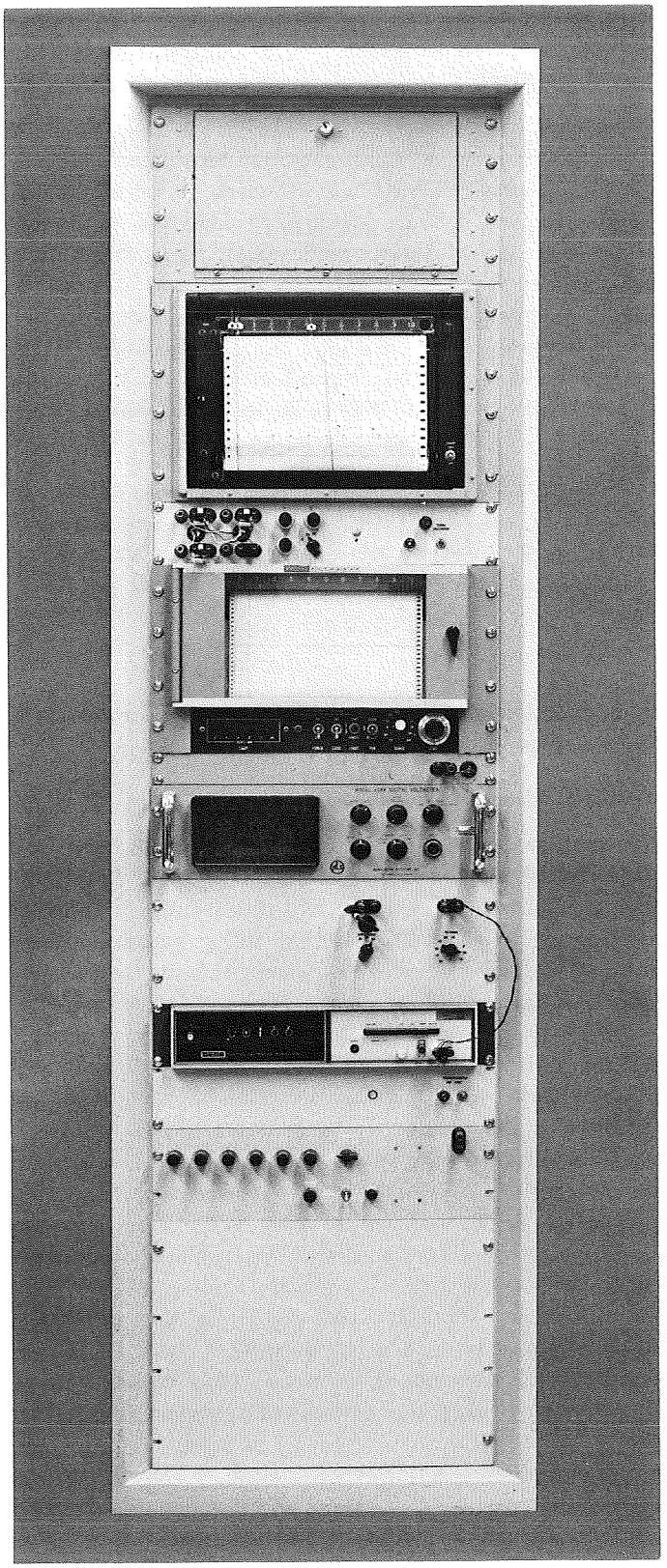


Fig. 3. Instrumentation rack with strip chart recorders and digital voltmeters

noisy performance has impaired accuracy. Therefore, new strip chart recorders have been ordered from Instrument Loan Pool.

The calorimeter and vacuum system have been checked out. Some initial tests have been made to verify guard design adequacy and to demonstrate evidence of degradation from short overlaps and seams. Some difficulties have been encountered with noisy recordings and shield tension due to thermal shrinkage. Improvements are being incorporated to eliminate these difficulties. With careful shield installation, reproducibility and low effective emissivity values (0.002) seem to be attainable.

Tests are planned to cover effects due to a variety of overlaps, seams, local compressions, and penetrations. Various factors will be investigated, such as attachment techniques, size of defect, seam tension, and patching effects.

B. A Preliminary Study of Nonlinearly Coupled Structural Systems, *J. Chen*

1. Introduction

In more recent years, several methods have been devised by which the dynamic behavior of a structural system may be predicted in terms of the properties of its component subsystems. In all of these, supplementary coupling conditions expressing compatibility and equilibrium at the coupling points of the subsystems must be prescribed. If the dynamic characteristics (e.g., receptance) of the component subsystems are determined experimentally, the mere prescription of the coupling condition is insufficient. In most cases the coupling of two subsystems at the interconnection points, at which the receptances were measured, introduces between the subsystems additional action that more than satisfies the coupling conditions. In Ref. 1, this has been accounted for by the introduction of so-called coupling units. These coupling units simulate mathematically the deformational behavior of such items as springs, bolts, washers, dampers, air pumping in lapjoints, etc., which cannot be included during the experimental determination of subsystem receptances. However, in many cases, these coupling units cannot be considered "linear," i.e., the relative motion of coupling units is not linearly proportional to its transmitted force. In other cases subsystems are interconnected by coupling units designed to achieve specifically prescribed coupling effects. A typical example of the latter is the solar panel system of the *Mariner* spacecraft, where four plate-beam-type structures are interconnected by eight highly nonlinear dampers.[1]

As in Ref. 1, this study is based on the concept of receptance and assumes that the subsystems are linear and have linear but not necessarily small damping. However, the coupling units here are assumed to be nonlinear. The general objective of this study is, then, to analytically and experimentally investigate the nonlinearly coupled subsystems.

2. General Formulation

The derivation of the receptance equation is detailed in Ref. 1, and will not be repeated here. However, the results will be briefly described. The response of the uncoupled system can be expressed as

$$\{D\} = [\overline{\overline{D}}]\,\{M\} + [\overline{D}]\,\{P\} \tag{1}$$

where $\{D\}$ is the frequency response vector of the structural system at the coupling points, $[\overline{\overline{D}}]$ is the receptance matrix between coupling points, $[\overline{D}]$ is the receptance matrix between coupling points and points of external excitation, $\{M\}$ is the constrained force matrix transmitted by the coupling units, and $\{P\}$ is the external excitation vector.

In a linear system, the constrained force is proportional to the relative motion of the coupling points. Since the coupling unit is in general elastic as well as dissipative, the proportionality constant is generally complex and frequency dependent. It can be formulated as

$$[\sim k \_]\,[C]\,\{D\} = \{\overline{M}\} \tag{2}$$

where $[\sim k \_]$ is the diagonal stiffness matrix of the coupling units, the elements of which are the inverse of the constraint matrix in Ref. 1; $[C]$ is a coupling matrix and $\{M\}$ is the column matrix, which includes the corresponding constrained forces. It can be verified that

$$[C]^T\{\overline{M}\} = \{M\} \tag{3}$$

Equations (1), (2), and (3) give the coupling constrained force

$$\{\overline{M}\} = \left[[\sim k \_]^{-1} - [C]\,[\overline{\overline{D}}]\,[C]^T\right]^{-1}[C]\,[\overline{D}]\,\{P\} \tag{4}$$

[1]Chen, J. C., *Mariner '69 Spacecraft Solar Panel Damper Analysis and Test Report*, Mar. 1969 (JPL internal document).

If the external excitations vanish, the nontrivial solution exists only if

$$\left|[\smile k \smile]^{-1} - [C]\,[\overline{\overline{D}}]\,[C]^T\right| = 0 \tag{5}$$

Equation (5) gives the natural frequencies (free vibration) of the entire coupled system if there is no damping. Then the corresponding modal response at the coupling points can be calculated by Eqs. (2) and (4).

Two examples are given.

(1) *Mass–spring system.* Figure 4 shows the mass–spring system and its subsystems. D_1, D_2 and M_1, M_2 are the displacements and constrained forces, respectively, and the receptance matrix $\{\overline{\overline{D}}\}$ is

$$\{\overline{\overline{D}}\} = \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{bmatrix} \tag{6}$$

where

$$\alpha_{11} = \frac{1}{k_1 - m_1\omega^2}, \qquad \alpha_{22} = \frac{1}{k_2 - m_2\omega^2} \tag{7}$$

and k_1, k_2, k_3 are the spring constants as shown in Fig. 4. Upon substituting Eqs. (6) and (7) into Eq. (5), one obtains the frequency equation

$$\omega^4 - \omega^2\left(\frac{k_1 + k_3}{m_1} + \frac{k_2 + k_3}{m_2}\right) + \frac{k_1k_2 + k_2k_3 + k_3k_1}{m_1m_2} = 0 \tag{8}$$

Equation (8) is the well-known frequency equation of the mass–spring system shown in Fig. 4.

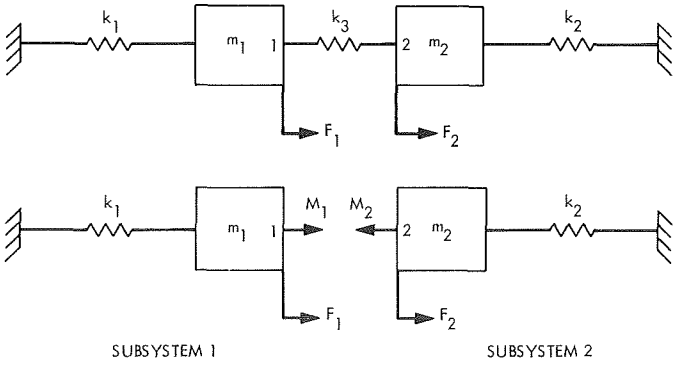


Fig. 4. Mass–spring system

(2) *Two uniform beams connected by a spring.* Figure 5 shows the entire structural system and its subsystems. Equation (5) can be written as

$$\left|\frac{1}{k} - [1 \quad -1]\begin{bmatrix} \alpha_{44} & 0 \\ 0 & \alpha_{55} \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right| = 0 \tag{9}$$

where α_{44} and α_{55} are the receptances:

$$\alpha_{44} = \alpha_{55} = -\frac{\cos\lambda L \sinh\lambda L - \sin\lambda L \cosh\lambda L}{EI\lambda^3(\cos\lambda L \cosh\lambda L - 1)} \tag{10}$$

where

$$\lambda^4 = \frac{\omega^2 A\rho}{EI}$$

and

A = cross-sectional area of beam

ρ = mass density of beam

EI = bending rigidity of beam

L = length of beam

ω = natural frequency of beam

Substituting Eq. (10) into Eq. (9), one obtains the frequency equation as

$$\frac{EI\lambda^3}{k}(\cos\lambda L \cosh\lambda L - 1) + (\cos\lambda L \sinh\lambda L - \sin\lambda L \sinh\lambda L) = 0 \tag{11}$$

This frequency equation has been verified by the classical method.

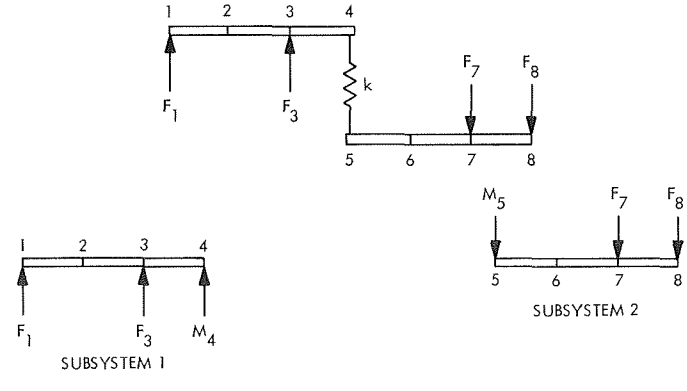


Fig. 5. Two-beam system

3. Some Observations on Nonlinear Coupling Units

If the subsystems are nonlinearly coupled, Eq. (2) is then no longer true. Before a true mathematical representation of a nonlinear coupling unit can be derived, one must understand the nature of the coupling unit and its nonlinear characteristics under various dynamic inputs.

As an example, we consider coupling units designed to achieve specifically prescribed coupling effects. The dampers used on the solar panel system of *Mariner* spacecraft consist of two spring-centered concentric tubes with the annulus filled with silicone fluid which is sealed by a pair of O-rings. The dampers have been dynamically tested and some of the results will be discussed here. The detailed test procedure is given in Footnote 1.

Figure 6 indicates the damper test setup and its instrumentation. The damper is fixed at one end through a load cell, i.e., $F_2(t)$, and the other end is connected to a shaker. The shaker oscillates periodically with constant maximum amplitude.

Figure 7a shows a picture taken from the oscilloscope; the vertical axis is the damping force, which is obtained by subtracting the spring force kx (where k is the "normal" spring constant and x is the damper displacement) from the total force, $F_2(t)$. The horizontal axis is the damper velocity. It shows that the damping force is not linearly related to the damper velocity. The damping force is also a function of frequency and amplitude. The

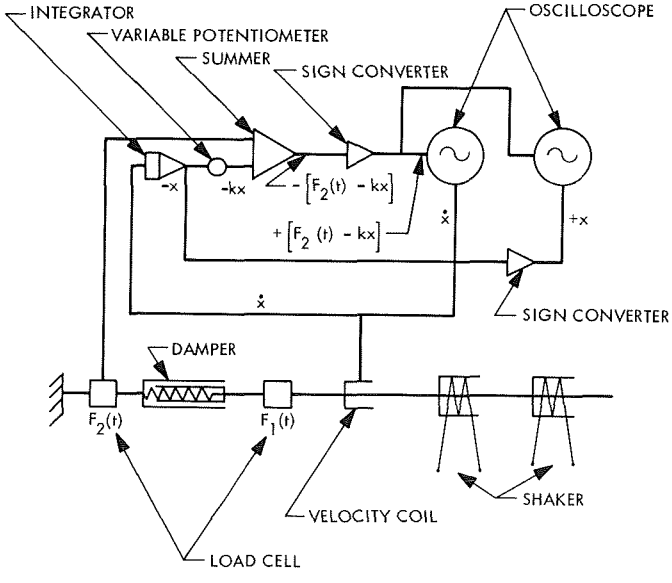


Fig. 6. Damper test setup

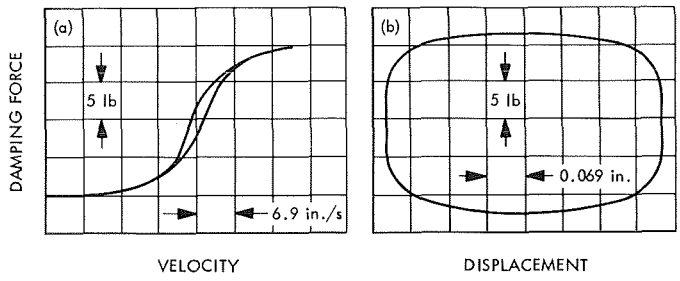


Fig. 7. Lissajous curves of damping force and (a) velocity and (b) displacement

damping force is approximately independent of the displacement; however, the damping forces are different for the same damper displacement if the frequencies and maximum amplitudes are different, as shown in Fig. 7b.

This shows that the damper is a highly nonlinear coupling unit and the damping force is nonlinear with respect to both velocity and maximum damper amplitude. However, a precise mathematical representation of the damper is still a question.

A different approach to the problem has been made. Let the damper oscillate at a certain frequency and maximum displacement; take the signal from load cell $F_2(t)$ through a very narrow band filter. If the narrow band filter sweeps from zero frequency to a very high frequency, the filtered signal represents the harmonics of $F_2(t)$. In other words, if the damper is oscillating at a frequency ω and maximum displacement D, the damping force $F_2(t)$ can be expressed as

$$F_2(t) = D_0 \cos \omega_0 t + D_1 \cos \omega_1 t + D_2 \cos \omega_2 t + \cdots \tag{12}$$

(for simplicity, only cosine waves are considered), where $D_0, D_1, D_3, \cdots$, are functions of the damper maximum displacement D and its frequency.

Figure 8 shows typical harmonics of a damper oscillating at 20 Hz and 0.1-in. maximum displacement. Here the damping force $F_2(t)$ can be expressed as

$$F_2(t) = D_0 \cos \omega t + D_1 \cos 3\omega t + D_2 \cos 5\omega t + \cdots \tag{13}$$

where $D_0, D_1, D_2, \cdots$, can be measured from the signals. It is very interesting that only the odd harmonics are

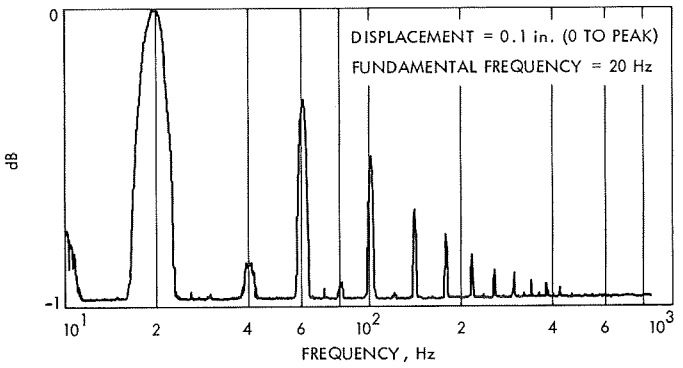


Fig. 8. Typical harmonic content of damper

strongly shown. Equation (13), at least, provides one possible mathematical representation of a nonlinear damper. Its usefulness will be discussed later.

It was intended to find a mathematical representation for the damper from the results of the damper test. The test results indicate that damping force is a function of displacement, velocity, frequency, and maximum amplitude of the periodic motion. However, one still cannot make a mathematical model with confidence that all the nonlinear effects found from the test will be included. Since our objective is to investigate the nonlinear coupling of linear systems, instead of looking for the mathematical representation of the coupling unit, one may ask what is the effect on the coupled system due to the nonlinearities of the coupling units. This is the approach that this study will follow.

4. Investigation of a Simple Nonlinearly Coupled System

Having discussed an actual nonlinear coupling unit, it seems proper that a mathematical attempt to investigate the nonlinearly coupled dynamic system should be made. A simple case will be considered; namely, two linear spring–mass systems coupled by a nonlinear spring. A similar linear system was given as an example in *Subsection 2.*

One of the commonest methods for treating nonlinear problems in mechanics is the perturbation method, which consists of developing the desired quantities in powers of some parameter that can be considered small, and determining the coefficients of the developments stepwise, usually by solving a sequence of linear problems. This technique has been successfully applied to the differential equations that govern the vibration system. The feasibility of applying a similar technique to the receptance method described in *Subsection 2* will be shown.

The system is shown in Fig. 4 except the spring k_3 is Duffing's type nonlinear spring. The receptance equation for the entire uncoupled system is given by Eq. (1); it can be written in detail as

$$\begin{Bmatrix} D_1 \\ D_2 \end{Bmatrix} = \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{bmatrix} \begin{Bmatrix} M_1 \\ M_2 \end{Bmatrix} + \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{bmatrix} \begin{Bmatrix} F_1 \\ F_2 \end{Bmatrix} \tag{14}$$

The constrained force transmitted by this nonlinear coupling unit is

$$\bar{M} = k_3 (D_1 - D_2) + \epsilon (D_1 - D_2)^3 \tag{15}$$

where $\epsilon \ll k_3$. From Eq. (3), the constrained force matrix $\{M\}$ can be obtained.

In Eq. (14), D_1, D_2 are the responses of the system and α_{11}, α_{22} are the receptances of the linear subsystems. For a harmonic motion, they are functions of frequency only. The nonlinearity enters only through the constrained force transmitted by the nonlinear coupling unit as shown in Eq. (15). If the excitation to the entire system is harmonic, one may make a reasonable assumption that the responses are also harmonic, even at the points where the nonlinear coupling units are attached. Then one may write

$$\begin{Bmatrix} F_1 \\ F_2 \end{Bmatrix} = \begin{Bmatrix} f_1 \\ f_2 \end{Bmatrix} \cos \omega t \tag{16}$$

and the response in the form

$$\begin{Bmatrix} D_1 \\ D_2 \end{Bmatrix} = \begin{Bmatrix} D_1^0 \\ D_2^0 \end{Bmatrix} + \epsilon \begin{Bmatrix} D_1^1 \\ D_2^1 \end{Bmatrix} + \epsilon^2 \begin{Bmatrix} D_1^2 \\ D_2^2 \end{Bmatrix} + \cdots \tag{17}$$

Substituting Eqs. (15), (16), and (17) into Eq. (14), and equating terms with like order of ϵ, one can then solve each equation stepwise and combine them to obtain the final solution. After the mathematical manipulation, the nonlinear response up to the order of ϵ can be expressed as

$$\begin{Bmatrix} D_1 \\ D_2 \end{Bmatrix} = [G] \begin{Bmatrix} \cos \omega t \\ \cos 3\omega t \end{Bmatrix} + 0(\epsilon^2) \tag{18}$$

where the matrix $[G]$ is a function of α_{11}, α_{22}, f_1, f_2, k_3, and ϵ. The convergence of the solution requires that

$$\begin{Bmatrix} D_1^1 \\ D_2^1 \end{Bmatrix} < 0\left(\frac{1}{\epsilon}\right) \tag{19}$$

Another method to solve the weak nonlinear problems is the so-called "slowly varying parameters" method, in which the response is assumed to be of the form

$$\begin{Bmatrix} D_1 \\ D_2 \end{Bmatrix} = \begin{Bmatrix} A_1(t) \\ A_2(t) \end{Bmatrix} \cos \omega t \tag{20}$$

Substituting Eq. (20) into (14) and using Eqs. (15) and (16), one obtains

$$\begin{Bmatrix} A_1 \\ A_2 \end{Bmatrix} = \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{bmatrix} \left[\begin{bmatrix} -k_3 & k_3 \\ k_3 & -k_3 \end{bmatrix} \begin{Bmatrix} A_1 \\ A_2 \end{Bmatrix} + \epsilon \begin{Bmatrix} -(A_1 - A_2)^3 \\ (A_1 - A_2)^3 \end{Bmatrix} \cos^2 \omega t + \begin{Bmatrix} f_1 \\ f_2 \end{Bmatrix} \right] \tag{21}$$

Now, although $A_1(t)$ and $A_2(t)$ are functions of time, t, it is assumed they vary slowly with respect to time. Therefore, one may integrate Eq. (21) through a period, i.e., from 0 to 2π, and let

$$\overline{A}_{1,2} = \frac{1}{2\pi} \int_0^{2\pi} A_{1,2}(t)\, dt \tag{22}$$

Then, Eq. (21) becomes

$$\begin{Bmatrix} \overline{A}_1 \\ \overline{A}_2 \end{Bmatrix} = \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{bmatrix} \left[\begin{bmatrix} -k_3 & k_3 \\ k_3 & -k_3 \end{bmatrix} \begin{Bmatrix} \overline{A}_1 \\ \overline{A}_2 \end{Bmatrix} + \frac{\epsilon}{2} \begin{Bmatrix} -(\overline{A}_1 - \overline{A}_2)^3 \\ (\overline{A}_1 - \overline{A}_2)^3 \end{Bmatrix} + \begin{Bmatrix} f_1 \\ f_2 \end{Bmatrix} \right] \tag{23}$$

This equation will give $\overline{A}_1$ and $\overline{A}_2$ in terms of receptance which is a function of frequency. Therefore, Eq. (23) is the frequency response equation.

If we let

$$m_2 = \infty$$

$$k_1 = 0$$

$$A_2 = f_2 = 0$$

$$\alpha_{22} = 0$$

$$\alpha_{11} = -\frac{1}{m_1 \omega^2}$$

the system reduces to a single-degree-of-freedom system; Eq. (23) becomes

$$\left(\omega^2 - \frac{k_3}{m}\right) \overline{A}_1 + \frac{f}{m} - \frac{\epsilon}{2m} \overline{A}_1 = 0 \tag{24}$$

This is the well-known frequency response equation of Duffing's nonlinear mass–spring system.

5. Concluding Remarks

A real nonlinear coupling unit used on a spacecraft and its experimentally determined characteristics have been presented. Also, a simple example of a nonlinearly coupled system and its analysis was given. Experiences have shown that the nonlinear effects on an entire structural system due to the nonlinear coupling units cannot be neglected. The purpose of this research project will be not only to develop a consistent technique to solve these kinds of problems, but also to investigate the nonlinear effects of a complex structural system if nonlinear coupling units are presented.

Transient loading to the nonlinearly coupled systems will also be investigated, particularly the time varying frequency sweep forcing function which is often imposed as a structural test. Analysis by the receptance method is in such a form that this varying frequency forcing function can be incorporated with relative ease. This will also be included in the future when the general theory on nonlinearly coupled dynamic system is developed.

Reference

1. Heer, E., "Coupled Systems Subjected to Determinate and Random Input," *Int. J. Solids Structures*, Vol. 3, pp. 155–166, 1967.

C. The *Mariner* Mars 1969 Temperature Control Flux Monitor, J. A. Plamondon

1. Introduction

The temperaure control flux monitor (TCFM) is a device used to measure integrated thermal radiation. The instrument was included as an engineering experiment on *Mariners VI* and *VII* with the following objectives:

(1) Provide a comparison of the thermal performance of a spacecraft in a simulated test environment against its performance in space where the reference for comparison is incident solar flux.

(2) Provide a direct measurement of the absolute value of the solar constant in space.

(3) As a result of objectives (1) and (2), establish a standard for future simulator testing.

2. Functional Description

The TCFM is a device used for measuring total radiance that requires only a general knowledge of the spectral radiance of the source. The measurements are absolute in the sense that calibration against a radiometric standard source is not required. An absolute error of less than $\pm 1.5\%$ was set as a design goal. A probable measurement error of about $\pm 1.0\%$ was achieved; however, its achievement required an in-flight calibration which was not part of the preflight planning. The resolution of the instrument is ± 250 $\mu W/cm^2$, or approximately $\pm 0.2\%$ of measured value at earth and $\pm 0.4\%$ of measured value at Mars. Because of the mechanization of the instrument, data having a relative resolution error as low as $\pm 0.03\%$ or ± 40 $\mu W/cm^2$ can be obtained.

The TCFM consists of a transducer and two trays of electronics (Fig. 9). The cylindrically shaped transducer contains a conical cavity for the sensing or active element of the transducer, which can be seen as a black opening in the face of the transducer. The purpose of the electronics is to control the transducer, and to interface with other spacecraft subsystems, in particular, power and telemetry. The electronics is almost entirely digital. The transducer is located on the spacecraft so as to provide an unobstructed view of the sun.

Figure 10 is a simplified cross-sectional view of the transducer showing its basic elements. As can be seen, the transducer consists of a conical shell, blackened on its interior surface and surrounded by a cylindrical-shaped thermal guard. The optical coating for the conical receptor is Parson's optical black lacquer. Also located within and on the guard structure and on the exterior of the conical shell are a number of fine wire windings and platinum thermometers. The wire windings and thermometers serve as electrical heater and temperature sensing elements for the control of the transducer. The platinum thermometers form the active elements of three bridge networks. Reference elements for the bridges are also located within the guard and are wire-wound.

The electronics uses these bridges in the following manner. A bridge located in the guard structure is so trimmed that when nulled by electrical heating, the guard settles at a precisely known, preselected temperature which is approximately 136°C. This temperature is controlled to ± 0.1°C. A second bridge having one platinum element in the guard and another on the cone is also so trimmed that, when nulled, the temperature of the cone is precisely slaved to the temperature of the guard to within ± 0.1°C. The third bridge is used to measure independently the temperature limit cycling of the guard structure within the ± 0.1°C deadband of the guard servo. This vernier measurement is relative and is accurate to ± 0.007°C with a tolerance of about 10%. The required electrical heating for nulling of the guard bridge and cone-guard bridge is supplied as a series of rectangular pulses of constant voltage which occur regularly in time. The pulse width is variable in discrete or digital steps called bits from 0 to 1023. Each discrete step adds or subtracts approximately 250 μW to the rectangular pulse.

This mechanization results in a simple energy balance equation for calculating an unknown input flux:

$$S = \frac{\epsilon}{\alpha}\sigma T^4 - \frac{P}{\alpha A}N \qquad (1)$$

where

S = solar flux (power/area)

ϵ = effective emittance

α = effective absorptance

σ = Stefan–Boltzmann constant

T = set point temperature

P = power/bit

N = number of bits

A = aperture area

Since the temperatures of both cone and guard are the same, no net heat transfer occurs between the two; consequently, all net radiant transfer must occur through the aperture. With the operating temperature preselected and known, the total output can be calculated from σT^4. The unknown input can in turn be calculated by subtracting from that output the amount of electrical power needed to maintain the cone temperature. Of course, the optical properties of the cone must be known and inserted into this equation properly.

From Eq. (1), it can be seen that the only calibrations needed are the set point temperature, the aperture area, and the amount of electrical power in a single power bit.

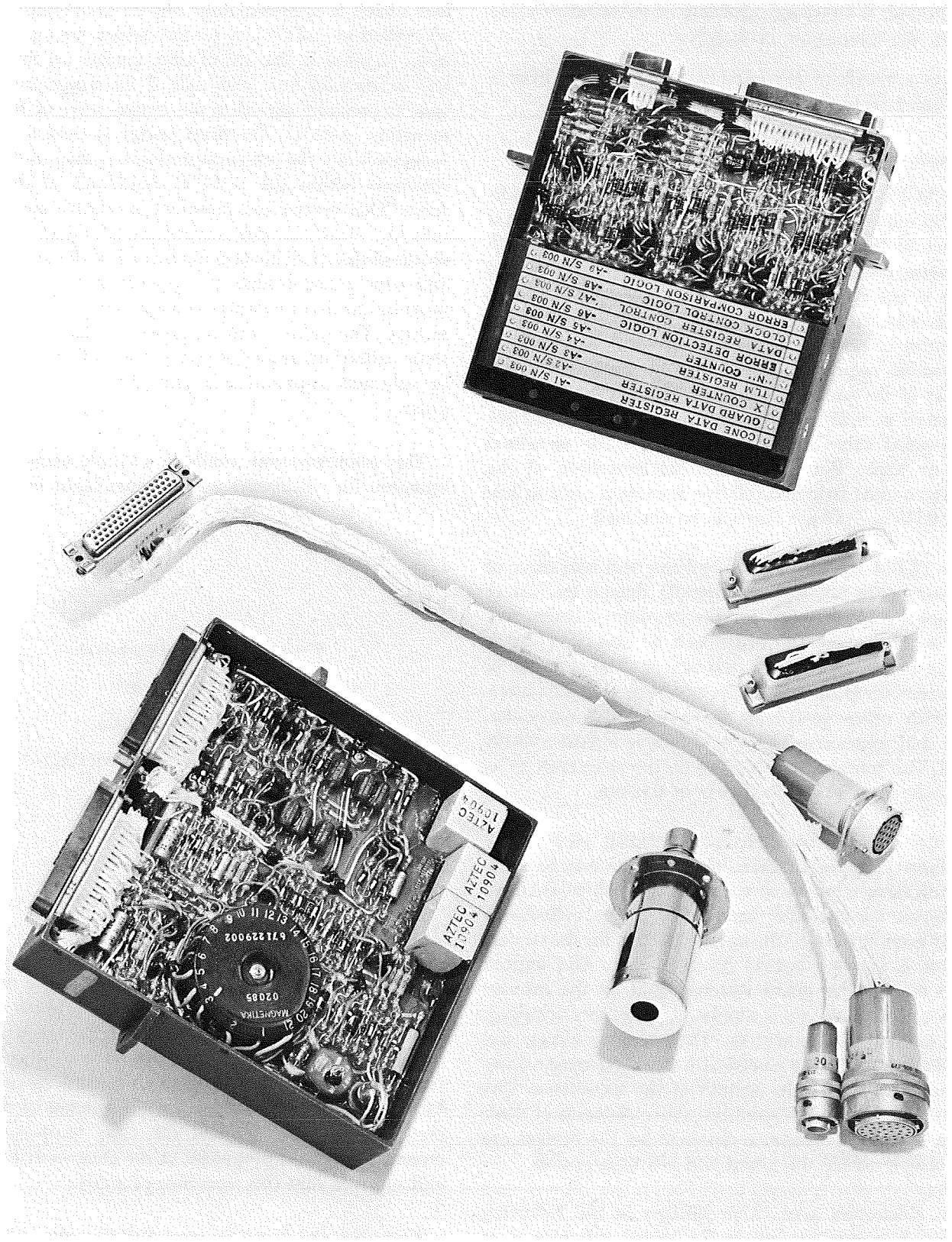


Fig. 9. TCFM components

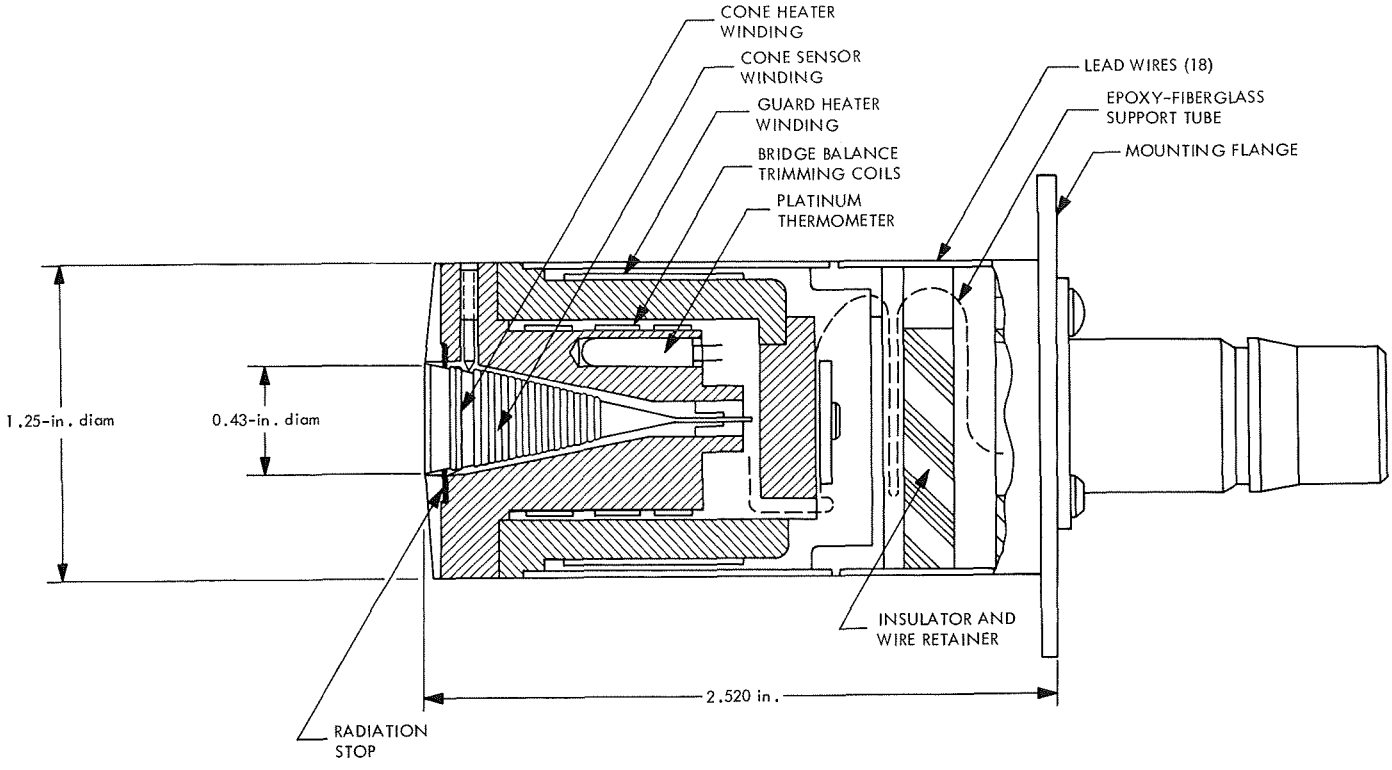


Fig. 10. TCFM transducer

With these quantities calibrated and the optical properties known, the remaining variable quantity is the width of the power pulse to the cone heater, i.e., the number of power bits. It is this number that is telemetered as a 10-bit binary word. Also telemetered is the cone-guard servo loop error for a data valid indication and the oscillation of the guard temperature within the guard servo deadband.

3. Flight Results and Discussion

Figure 11 shows the solar flux plotted against the heliocentric range for both *Mariners VI* and *VII*. Note that fairly good agreement was obtained between the two instruments. Figure 12 shows the same data normalized or corrected to 1 AU and plotted against GMT in days. This permits a comparison of data taken at corresponding times. Note that the flux scale is highly expanded, and there is a negative slope which is large in relationship to the $\pm 1.5\%$ error limit. The negative drop is approximately 5%. For a constant solar constant and stable instrument, the data should result in a horizontal plot that is at least consistent with the $\pm 1.5\%$ error limit.

The humps and dips are real to a relative resolution of about 40 $\mu W/cm^2$ or less, or about 0.03% of value. A comparison of the two sets of data shows the same general behavior wherein what appears to be random flux level changes of typically 0.1% or so were detected at corresponding times. However, exact time correlation is not possible because of insufficient simultaneous tracking. Each spacecraft was tracked no more than 30–40% during the flight. This plot, combined with other data, reveals no inconsistency in these small flux level changes between the two sets of data. Though good correlation exists in regard to changes in flux level at corresponding times, the absolute values of the fluctuations typically do not agree. This is to be expected due to the digital mechanization of the instrument.

As a result of the data plotted here, complemented with the total accumulation of lower resolution data, the following observations can be made:

(1) Variations in solar flux of the order of tenths of 1% do occur and, as far as can be determined, these variations appear to be random. To this time, no serious attempt has been made to correlate with solar activity, such as flares and sunspots.

(2) The TCFM sampled every 0.2 s and its data was telemetered every 7 min. Because of this, it appears

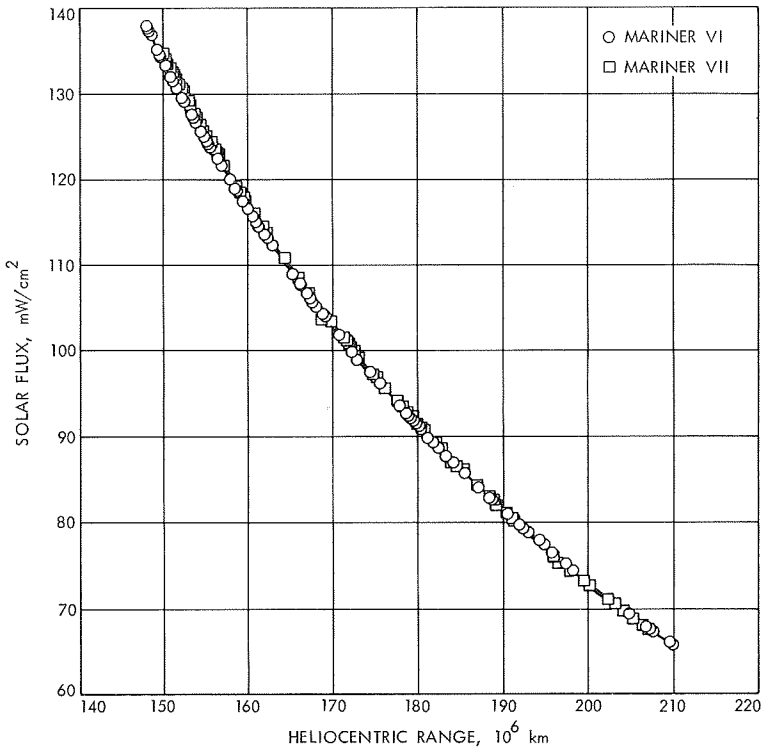


Fig. 11. Solar flux vs heliocentric range for *Mariners VI* and *VII* (launch through encounter)

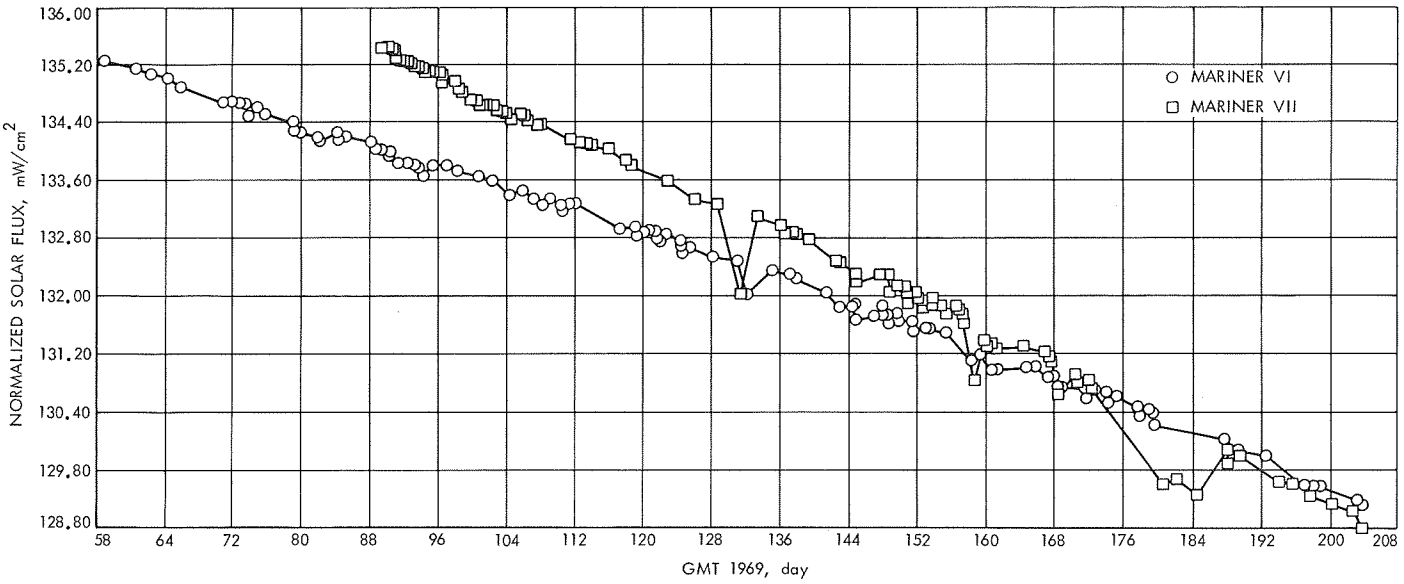


Fig. 12. Solar flux (normalized to 1 AU) vs GMT for *Mariners VI* and *VII* (launch through encounter)

that no variation larger than ±0.2% had occurred on this time scale, since no variation had been observed that was sufficiently large to drive the error detector to update in two successive telemetry frames.

(3) Those variations that occur, occur over periods of hours with the shortest observed period being roughly 2 h.

(4) The largest observed variation was 0.4%.

(5) Longer term cyclic or oscillatory variations of several percent over a period of days were *not* observed.

Such long term oscillatory variations have been reported in the literature by reputable sources. These sources report variations of up to 3% over 10-day periods and present evidence of other periodic behavior of a complex nature.

The general downward slope is now considered. After observing the decreasing trend for some time, numerous causes were postulated. After investigating each possibility, most were eliminated with the following exceptions:

(1) The results were real and the absolute difference between the *Mariner VI* and *VII* data at a given time could be attributed to calibration offset.

(2) A drift in transducer set point or in power per bit, or both, was occurring.

(3) The optical coating of the receptor was degrading (bleaching) at a rate far in excess of what was anticipated.

Of the three, the least probable or least believable was degradation in optical properties. Extensive testing of the stability of the coating was conducted, wherein samples of the coating were subjected to long-term ultraviolet degradation in vacuum and at operating temperature. These tests indicated that at worst 0.2% to 0.4% changes in absorptance might occur and that such changes would be exponential over typically a 1000-h exposure period. Additionally, no measurable change in emittance was observed. The above degradation values would not nearly account for the decreasing trend, which required an order of magnitude greater degradation spanning much more than the 1000 h. However, the possibility of degradation still cannot be completely dismissed since the space environment is far more complex than our simple UV environment, and the interaction of the space environment with materials is really not well understood.

The drift explanation was most plausible since drift in the set point temperature of the transducer had been experienced in laboratory testing. Five complete units and one additional spare transducer were built. In testing these units, including the spare transducer, a drift amounting to 1.5% had been observed over comparable periods for all transducers. Repeated testing of each transducer led us to believe that the drift was associated with aging of the wire-wound cone-guard bridge, and that the aging problem was short term, being essentially complete in the first 500 h of operation. However, it was never possible to obtain a solid relationship for drift versus time in these tests since a large portion of the operating time was logged in uncontrolled environments. This drift was considered essentially complete at launch by which time each transducer had been subjected to at least 1000 h of operational testing. During the combined 6000 to 7000 h of logged operational time, no drift had been recorded in power per bit for any instrument.

At this point we were reluctantly willing to accept the drift explanation, with some reservations regarding degradation, except that some confirming results were presented to us from other measurements from the spacecraft. In mid-June, the reference solar cell and the solar panel temperature data suggested a possible change in solar flux. In both cases the data indicated that incident solar flux was falling off at a rate faster than $1/r^2$, and indeed the data seemed to be in fair agreement with the TCFM. Then, just before the encounter of *Mariner VI*, additional information came to our attention from the trajectory area. In particular, residual errors in trajectory calculations, if lumped into solar pressure, produced results similar to the TCFM data. However, this data could not be taken too seriously since it was well known that some degradation in the optical properties of the solar panels occurs, and such degradation could possibly explain all three effects. Nevertheless, the data was there and if it could be confirmed, the results would have been exciting and unexpected. A varying solar constant of this magnitude would raise some serious questions, and could have explained why so many values for the constant exist.

To resolve the problem, an independent in-flight calibration of the TCFM was needed. For this purpose, the only available reference was black sky which, in a practical sense, was a zero flux reference to the TCFM. However, because the TCFM has a near-hemispherical

response, a slightly greater than 90-deg pitch maneuver was required. Such a maneuver was proposed, but because of risk to the spacecraft, it was not permitted until after encounter.

After encounter a pitch turn of about 100-deg was performed with *Mariner VI*. The calibration obtained showed a significant shift in the zero of the instrument. Using this calibration point and assuming that all of the drift had occurred in the set point temperature of the instrument, a new temperature was computed and an exponential equation for set point temperature versus time was fitted to the initial, prelaunch, and post-encounter calibrations. The resulting function was then used to reduce the *Mariner VI* data gain. The results are given in Fig. 13, which shows normalized solar flux as a function of heliocentric range. As can be seen, a nice horizontal plot results giving an average solar constant of 135.25 mW/cm^2. The only problems that remain to be resolved are the conflict imposed by the solar pressure, reference solar cell, and solar panel temperature data, and to place a final tolerance on the accuracy of the average solar constant value obtained. The first problem can be resolved in a purely TCFM-related way; namely, a pitch turn calibration of roughly 45 deg with *Mariner VI* at some future time to confirm the present drift function and, if possible, a pitch of at least 35 deg with *Mariner VII* to independently calibrate that instrument.

In regard to the second problem, if it is assumed that the solar flux is on the average constant, the fact that this plot is almost perfectly horizontal suggests that the various terms making up the coefficients of the data reduction equation have either very small residual errors or existing errors are self-compensating, with the exception of effective absorptance. This conclusion results from a detailed error analysis of the equation which will not be developed here. A residual error in the equation, except for absorptance, results in an unavoidable bow in this type of plot. An error in absorptance is an exception since it enters the flux equation only as a scaling factor. Assuming a constant solar flux, the absolute tolerance that can be assigned to the average value is +1% and −0.3%. Further tightening of these tolerances requires further knowledge of the absorptance and reflectance characteristics of Parson's black. In particular, we need to know more in regard to its directional reflectance characteristics and to obtain a better absolute value for absorptance. Such a study is currently under way.

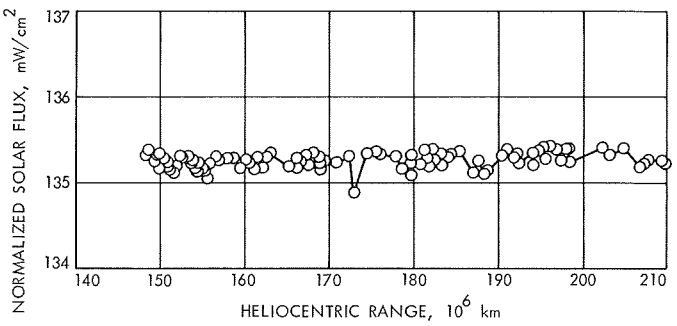


Fig. 13. Normalized solar flux vs heliocentric range for *Mariner VI* (corrected for drift)

Finally, in regard to our objectives as an engineering experiment, the comparison between simulator tests and flight results is good. In general, agreement was within 5°F with few exceptions. Therefore, the general conclusion is that the environment provided by the present 10-ft facility is adequate for high-quality temperature/vacuum testing of spacecraft whose design employs temperature control concepts as used in the *Mariner* Mars 1969. As for establishing a radiometric standard for total irradiance measurements, this has been accomplished. The techniques of calibration are simple and reproducible, and instruments upon which the TCFM was modeled are readily available.

D. Nonstationary Random Excitation and Response of Coupled Systems,

E. Heer and J. N. Yang

In-flight measurements taken during the launch and exit phase of a number of recent space flights show that the major spacecraft excitations and responses during any one flight consist of a series of highly transient pulses resulting from flight events such as booster engine cutoff, booster ignition, etc. An example of such transients is given in Ref. 1. The deterministic prediction of the time and spectral characteristics of any one of these events is associated with considerable uncertainties and statistical variations when going from one spacecraft or flight to another. To establish evaluation and design criteria for future spacecraft, it is therefore necessary to invoke probabilistic techniques and to allow for convenient updating of such criteria as new measurements become available. The receptance coupling program (RECEP),[2] a computer program, enables, at the present

[2]SPS 37-49, Vol. III, pp. 159–161, and SPS 37-56, Vol. III, pp. 164–165.

time, the treatment of stationary random processes; it is the objective of this work to extend the capabilities of RECEP so that nonstationary random phenomena can be treated as well. This will establish the capabilities alluded to above. The present article gives the necessary theoretical background for this development where it is presupposed that the system receptance matrix is available from RECEP.

Let $x(t)$ be the vehicle response vector due to an arbitrary excitation vector $p(t)$. The following relations are then valid:

$$x(t) = \int_{-\infty}^{\infty} h(t-\tau)\, p(\tau)\, d\tau \tag{1}$$

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\, e^{-i\omega t}\, dt \tag{2}$$

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega) e^{i\omega t}\, d\omega \tag{3}$$

where $h(t)$ is the unit impulse response matrix and $H(\omega)$ is the receptance matrix.

Substituting Eq. (3) in Eq. (1) yields

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\, P(\omega)\, e^{i\omega t}\, d\omega \tag{4}$$

where

$$P(\omega) = \int_{-\infty}^{\infty} p(t)\, e^{-i\omega t}\, dt \tag{5}$$

is the Fourier transform of $p(t)$.

The mean value function $\bar{x}(t)$ at the time instant t is the average of the ensemble of all the possible time histories $x(t)$ taken at t. Taking this expectation on both sides of Eq. (4), one obtains

$$E[x(t)] = \bar{x}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\, \bar{P}(\omega)\, e^{i\omega t}\, d\omega \tag{6}$$

in which

$$\bar{P}(\omega) = E[P(\omega)] = \int_{-\infty}^{\infty} \bar{p}(t)\, e^{-i\omega t}\, dt \tag{7}$$

is the Fourier transform of the mean value or expectation of $p(t)$,

$$\bar{p}(t) = E[p(t)] \tag{8}$$

Taking expectation on both sides of Eq. (1), one obtains

$$\bar{x}(t) = \int_{-\infty}^{\infty} h(t-\tau)\, \bar{p}(\tau)\, d\tau \tag{9}$$

Let

$$R_{\underset{jk}{*}}^{JK}(t_1,t_2)$$

be the cross-covariance at time t_1 and t_2 of the nonstationary responses $x_j^J(t)$ and $x_k^K(t)$ of the coupled system at point j in component system J and at point k in component system K, and let

$$R_{\underset{lm}{p}}^{LM}(t_1,t_2)$$

be the cross-covariance at times t_1 and t_2 of the nonstationary excitations $p_l^L(t)$ and $p_m^M(t)$ of the coupled system at point l in component system L and at point m in component system M. By definition one then has

$$R_{\underset{jk}{*}}^{JK}(t_1,t_2) = E[\{x_j^J(t_1) - \bar{x}_j^J(t_1)\}\{x_k^K(t_2) - \bar{x}_k^K(t_2)\}] \tag{10}$$

$$R_{\underset{lm}{p}}^{LM}(t_1,t_2) = E[\{p_l^L(t_1) - \bar{p}_l^L(t_1)\}\{p_m^M(t_2) - \bar{p}_m^M(t_2)\}] \tag{11}$$

Furthermore, let

$$S_{\underset{jk}{*}}^{JK}(\omega_1,\omega_2)$$

be the generalized response spectral density at frequencies ω_1 and ω_2 between point j in component system J and point k in component system K, and let

$$S_{\underset{lm}{p}}^{LM}(\omega_1,\omega_2)$$

be the generalized excitation spectral density at frequencies ω_1 and ω_2 between point l in component system L and point m in component system M. The following

double Fourier transform relations then hold:

$$S^{JK}_{\substack{*\\ jk}}(\omega_1,\omega_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R^{JK}_{\substack{*\\ jk}}(t_1,t_2)\, e^{-i(\omega_1 t_1+\omega_2 t_2)}\, dt_1\, dt_2 \quad (12)$$

$$S^{LM}_{\substack{p\\ lm}}(\omega_1,\omega_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R^{LM}_{\substack{p\\ lm}}(t_1,t_2)\, e^{-i(\omega_1 t_1+\omega_2 t_2)}\, dt_1\, dt_2 \quad (13)$$

with the following inversions:

$$R^{JK}_{\substack{*\\ jk}}(t_1,t_2) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S^{JK}_{\substack{*\\ jk}}(\omega_1,\omega_2)\, e^{i(\omega_1 t_1+\omega_2 t_2)}\, d\omega_1 d\omega_2 \quad (14)$$

$$R^{LM}_{\substack{p\\ lm}}(t_1,t_2) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S^{LM}_{\substack{p\\ lm}}(\omega_1,\omega_2)\, e^{i(\omega_1 t_1+\omega_2 t_2)}\, d\omega_1\, d\omega_2 \quad (15)$$

The response at point j in component system J due to the excitations in all N component systems at time t_1 is

$$x^J_j(t_1) = \sum_{L=A}^{N}\sum_{l=1}^{l_p}\int_{-\infty}^{\infty} h^{JL}_{jl}(t_1-\tau)\, p^L_l(\tau)\, d\tau \quad (16)$$

Using Eq. (16) and forming the appropriate quantities and substituting in Eq. (10) yields, with Eq. (11) and changes of variables, the following expression:

$$R^{JK}_{\substack{*\\ jk}}(t_1,t_2) = \sum_{L=A}^{N}\sum_{M=A}^{N}\sum_{l=1}^{l_p}\sum_{m=1}^{m_p}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^{JL}_{jl}(\tau_1) \times R^{LM}_{\substack{p\\ lm}}(t_1-\tau_1, t_2-\tau_2)\, h^{KM}_{km}(\tau_2)\, d\tau_1\, d\tau_2 \quad (17)$$

Changing the variables in Eq. (15), substituting in Eq. (17), and using Eqs. (2) and (14) give the generalized response spectral density

$$S^{JK}_{\substack{*\\ jk}}(\omega_1,\omega_2) = \sum_{L=A}^{N}\sum_{M=A}^{N}\sum_{l=1}^{l_p}\sum_{m=1}^{m_p} H^{JL}_{jl}(\omega_1)\, S^{LM}_{\substack{p\\ lm}}(\omega_1,\omega_2)\, H^{KM}_{km}(\omega_2) \quad (18)$$

Substituting Eq. (18) in Eq. (14) gives the desired result. In Eq. (18) the first and the third right-hand terms are: the receptance between points j,l in component systems J,L and points k,m in component systems K,M evaluated at frequencies ω_1,ω_2. Equation (18) reduces to the last equation in SPS 37-49, Vol. III, p. 160, if the excitations are stationary random functions, in which case the cross-covariance in Eq. (17) is a function only of $t_2 - t_1 = \tau$.

The evaluations of Eq. (14) require double Fourier inversion techniques. A two-dimensional fast Fourier transform and inversion technique has been developed and is presented in Ref. 2.

At the present time no method exists to determine with any degree of confidence the covariance function of a nonstationary process from a single sample function (or even on a few numbers of sample functions), and no physical significance of practical use is known for the generalized spectral density. Therefore, usually one has to be satisfied with the covariance function and its associated generalized spectral density of a time domain

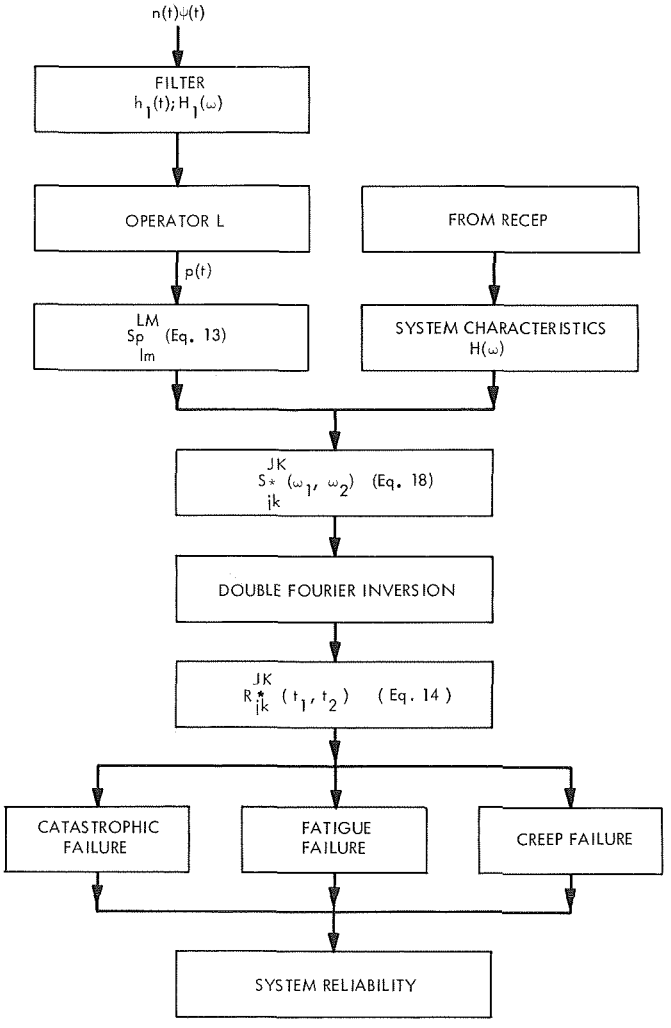
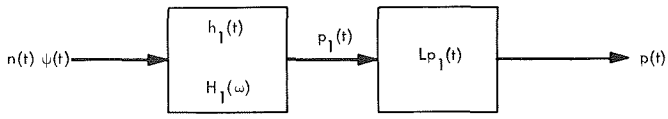


Fig. 14. Flow chart for development of subroutine computer program coupled to RECEP

"model" of the excitation process, which exhibits reasonably well a general trend of the variance function expected from the observed sample functions, and then reproduces from it a covariance function and its associated generalized spectral density within a certain time interval. For example, the following model has been used to



represent a severe gust loading or an earthquake excitation (Ref. 3) in which $n(t)$ is white noise, $\psi(t)$ is a deterministic function of time, $h_1(t)$ and $H_1(\omega)$ are, respectively, the unit impulse response function and the receptance of a filter. In other words, the input excitation is obtained by passing a non-stationary white noise through a filter and a differential (or integral) operator L. In the RECEP program, a general filter and a general operator (differential or integral) will be incorporated to generate the desired input $p(t)$.

The structure may fail due to three possible modes depending on the mission: (1) catastrophic failure, (2) fatigue failure, or (3) creep failure. The techniques of predicting both catastrophic failure and fatigue failure, using the mean value function and the cross-covariance matrix of the response process obtained previously, have been developed (Ref. 3)[3] while the prediction of creep failure is now under investigation.

Therefore, the RECEP program will contain the following packages: (1) general program for numerical double Fourier transform and inversion, (2) general program for generating the desired statistical characteristics of the excitation process to the system, and (3) program for predicting the system reliability.

A flow chart (Fig. 14) for the development of a subroutine computer program to be coupled with the present RECEP program has been designed and will be programmed as the next step of this work.

References

1. Heer, E., and Trubert, M. R., *Analysis of Space Vehicle Structures Using the Transfer-Function Concept,* Technical Report 32-1367. Jet Propulsion Laboratory, Pasadena, Calif., Apr. 1, 1969.
2. Yang, J. N., and Shinozuka, M., "Numerical Fourier Transform in Random Vibration," *J. Engineering Mechanics Division, ASCE,* Vol. 95, No. EM3, pp. 731–746, June 1969.
3. Shinozuka, M., and Yang, J. N., "On the Bound of First Excursion Probability," *J. Engineering Mechanics Division, ASCE,* Vol. 95, No. EM2, Apr. 1969. Also available as Technical Report 32-1304, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1968.

[3]Also Shinozuka, M., and Yang, J. N., "Nonstationary Peak Distribution," to be presented at the ASCE-EMD Specialty Conference on Probabilistic Concepts and Methods in Engineering, Purdue University, Nov. 12–14, 1969.

XV. Electronics Parts Engineering

ENGINEERING MECHANICS DIVISION

A. Accelerated Life Test Program for Monolithic Integrated Circuits, *M. F. Johnston*

1. Introduction

The problems in assessing the reliability of electronic parts are long duration and large-scale tests required to demonstrate low failure rates. Accelerated testing holds the possibility of reducing these problems; however, this technique is limited to the extent that the accelerated conditions can be related to performance at rated conditions.

A useful approach in designing an accelerated test is to run constant stress tests and step stress tests simultaneously. The constant stress groups can be run at several severity levels and the results used to plot a curve showing a measure of quality versus the measure of stress. The measure of quality may be failure rate, time to failure, etc.

The step stress level may be run using small, medium, or large steps at uniform time intervals. The size of the step determines the manner of assessing cumulative damage. Medium and small steps are analyzed accounting for cumulative damage: for medium steps by summation at each prior step, and for small steps by integration of previous cumulative damage.

A theory of cumulative damage is needed to relate the results of the various tests. The one usually chosen is the simple linear theory of the form of the Arrhenius equation:

$$rr = A \exp(-E/kT)$$

where

rr = reaction rate

A = a constant (also called frequency factor)

E = activation energy, eV/molecule

k = Boltzmann constant

T = temperature

2. Test Design

The test design included 8 groups of 30 Fairchild micrologic 9042 dual *nand* gates at the following life test conditions:

Step stress

Group	Time interval, h	Stress levels, V
21	72	21, 24, 27, 30, 33, 36, 39, 42, 44, 46, 49, 52, 55, 57, 60
22	120	
23	216	
24	336	

Constant stress

Group	Measured at time intervals of:	Stress levels, V
25	4, 12, 18, 29, 53, 99, 182, 336, 621, 1146 h	21, 27, 42, 49
26		
27		
28		

The test design included a preliminary studies phase wherein it was established that the device had two potential problem areas: (1) circuit susceptibility to leakage current, and (2) metallization. A test circuit was selected which would allow the highest possible current in the metallization and reverse bias on the transistor emitter–base junctions.

3. Test Results

The catastrophically failed units must be characterized as normal or abnormal in order to achieve a time to failure based on wear-out mechanisms. A broad definition of normal for these purposes shall be that the failure mechanism could be expected to occur in normal usage over an extended time interval of applied stress.

The test is approximately two-thirds complete and 40 devices which failed catastrophically have been analyzed. These results indicate that about 25% of the failures that have occurred thus far fall into the normal classification. The remainder of the failures all fall into one failure mode and mechanism classification and are the result of electrical overstress.

Analysis of the catastrophic failures for the step stress groups indicates that up to the tenth stress level the failures that have occurred are in the normal classification and have resulted from a legitimate acceleration of a wear-out mechanism. The stress level for the constant stress and step stress groups where abnormal failure mechanisms predominate was approximately the same.

The results indicate that the catastrophic failure up to a point is a function of accumulative stress rather than stress level only. The proposed models for accumulative degradation may be true for the final analysis with only slight modifications. This applies to the constant stress groups as well as the step stress groups.

The present testing is expected to be completed during the latter part of 1969. The validity of this accelerated test and the proposed accumulated stress model can then be determined.

XVI. Space Simulators and Facility Engineering

ENVIRONMENTAL SCIENCES DIVISION

A. Creep in Solid Metallic Mercury, *N. N. Nezhni*

1. Introduction

Information on solid metallic mercury at −95°F was needed to determine the design adequacy of the proposed tilting frozen mercury collector for the JPL Ion Engine Test Facility. A test program was initiated to gain the desired mercury data. The creep phenomenon encountered in the experiments was found to be a major restraint in the collector design.

2. Description of Test Program

The initial objective consisted of finding the following mechanical properties of mercury at −95°F: Young's modulus, shear stress, tensile stress, compressive stress, Poisson's ratio, adhesion, density, compressive slump, and creep.

In the absence of information on solid mercury, literature searches were conducted at JPL and the National Bureau of Standards at Boulder, Colorado. German and Russian documents were translated and personnel prominent in the cryogenic and mercury fields were contacted. Single crystal mercury data was found but nothing was uncovered that was directly applicable to the problem at hand.

A test program was then initiated to find the desired mercury data. Laboratory space, two small vacuum chambers, a Wild level, dry ice, and enough commercial grade mercury were obtained for the program. Sketches of jigs, molds, and fixtures were drawn and the parts fabricated. The mercury shapes were then cast.

Liquid mercury does not readily stick to most solid materials. However, at −100°F, it was found to be almost impossible to separate solid mercury (without deforming it) from molds made of low carbon steel, aluminum, stainless steel, plastic tubing, and pyrex. To overcome this adhesive problem, the solid mercury beams, rods, etc., used in these experiments were made in Teflon molds and mechanically pressed out (between frozen forms).

Young's modulus was to be calculated from the deflection data of the simply supported beam. The beam deflection test was set up as a flat beam with both ends simply supported with a centrally located weight. With load applied, the beam continually deflected. Upon load

removal, the beam deflection remained. The test was repeated and results were reproducible. The creep rate was such as to instigate the following tests:

(1) Compressive creep, 97.4 and 37.5 psi constantly applied.[1]

(2) Tensile creep, 42.6 psi constantly applied.[1]

(3) Beam bending creep.

(4) Static creep (with no external load).

3. Compressive Test

Figure 1 shows the configuration used for compressive creep tests. The figure also shows the type of deformation encountered with solid mercury at $-95 \pm 5°F$ when 97.4 psi (compression) was applied constantly. The data from compressive creep test 1 are plotted in Fig. 2 and show the creep in centimeters vs time in hours. It is a typical creep curve for ordinary metals held at constant temperature in a furnace (see Ref. 1). The curved portions of primary creep, secondary creep, and tertiary creep are very pronounced, and it is evident that the creep rate of mercury at $-95°F$ is very high. The curve in Fig. 3 conforms to the findings in Fig. 2 and is plotted as (computed) percent creep vs time.

In compressive creep test 1, the mercury collapsed to one side, binding the carriage bolt and top compression member. The test was repeated (compressive creep test 2) with side channel guides to help attain vertical compressive forces. In creep test 2, the mercury deformed from front to back. The data for creep test 2 (Figs. 2 and 3) duplicate those for creep test 1, confirming that 97.4 psi creates high creep rates for frozen mercury at $-95°F$.

For compressive creep test 3, the test jig was refabricated to prevent off-center compressive forces. The mercury bar used had a cross-sectional area of one square inch, and with 40 lb constantly applied (subtracting the weight of the carriage and mercury), the applied compressive stress was 37.5 psi. This creep test was conducted

[1]The loads applied in these tests relate to the static loading one obtains from a 1- × 1-in. (square) column of frozen mercury 80 in. high. The static load near the bottom of this theoretical column is 80 in. × 0.512 lb/in.^{3} = 41 psi. Therefore, a load of 97.4 psi has a safety factor of 2.37; other loads used were aimed at a safety factor in the vicinity of unity.

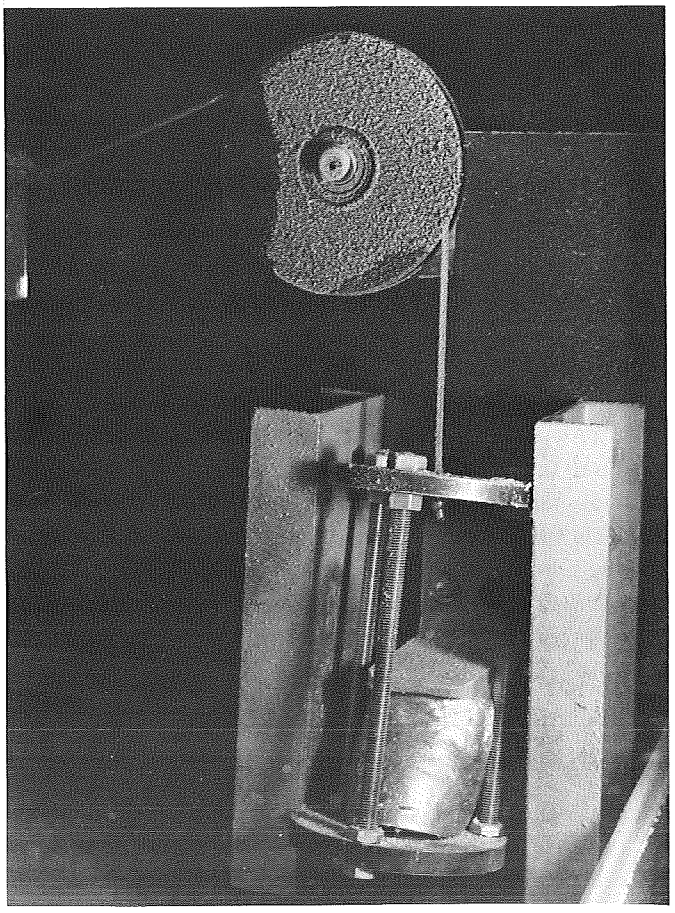


Fig. 1. Compressive creep test configuration

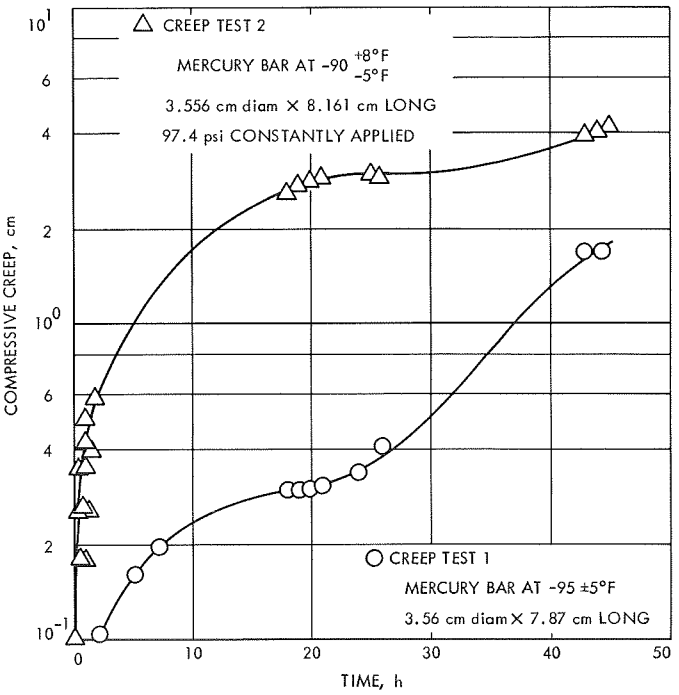


Fig. 2. Compressive creep tests 1 and 2

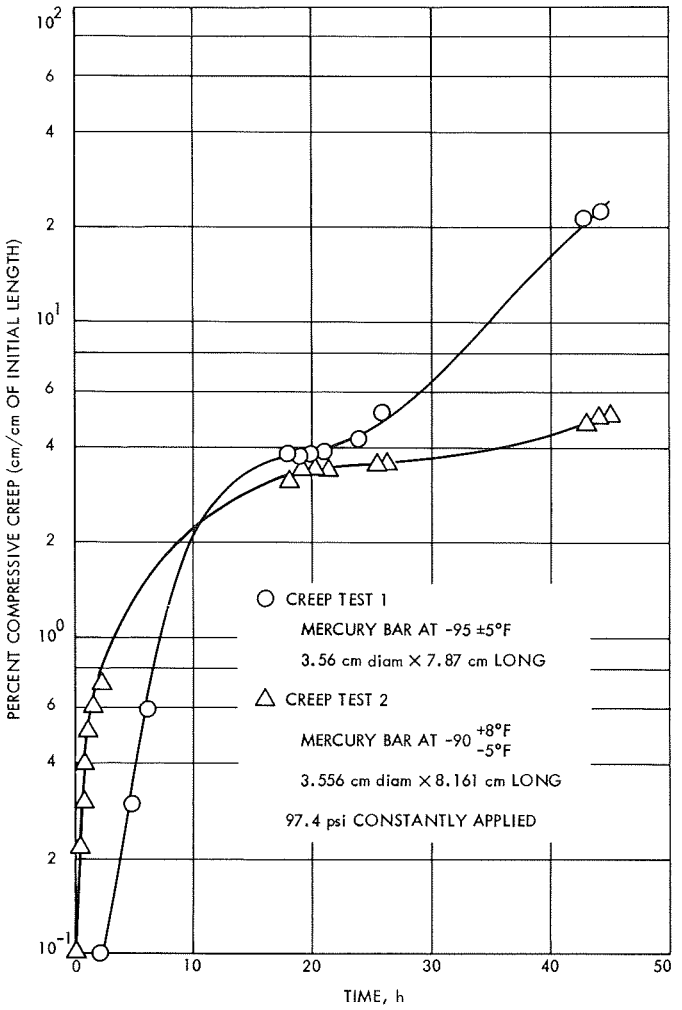


Fig. 3. Percent creep vs time

for over 190 h, and the results are plotted in Fig. 4. This curve resembles the previous plots and it also indicates that the compressive creep is large at 37.5 psi and −95°F.

4. Tensile Test

Figure 5 shows the comparison between the initial tensile test specimen (before tension was applied) and the two segments remaining after the break. The ruptured ends resemble the ends of pencils terminating in approximately 60-deg cones (total included angle). The test specimen width of 0.5 in. necked down (before the cone) to 0.315 in., and the 0.375-in. dimension constricted down to 0.225 in. The specimen linear elongation was 38% (0.90 in./2.375 in. × 100). The curve in Fig. 6 gives the tensile creep as a function of time for a constant load of 7.99 lb. This curve has the same characteristics as the others given for compressive creep, i.e., three pronounced

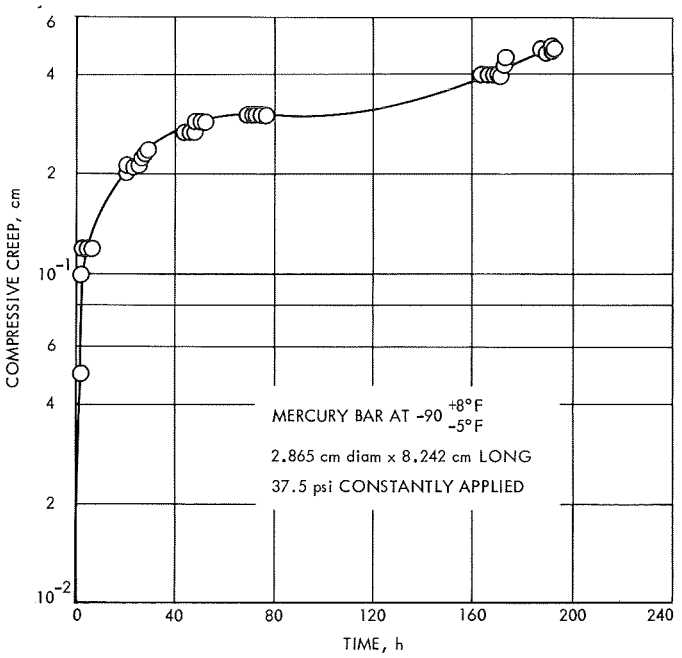


Fig. 4. Compressive creep test 3

time–creep periods. It is also evident that 42.6 psi tensile stress induces considerable creep in solid mercury at −95°F.

5. Beam Bending Test

Figure 7 shows two frozen mercury beam specimens, 0.25 × 1 × 8 in., with the ends simply supported. A 1500-g weight is slung from the center of the rear beam. The test data for the beam with the 1500-g weight is plotted in Fig. 8 and gives the characteristic creep curve. As the 1500-g weight on the solid mercury beam induced large creep rates, a 1000-g weight was applied to an identical bar (front beam, Fig. 7) to obtain additional data. It was found that the "knife edge" copper end supports (used in this test) induced lateral tensile forces making computation of bending stresses difficult. Therefore, this test is presently being rerun with Teflon rollers as edge supports so that the beam will be subjected to bending without lateral tension.

6. Static Test (No Externally Applied Load)

The lower right hand corner of Fig. 7 shows a 5- × 5- × 1.375-in.-thick slab of solid mercury held at −95°F for 1 mo. Horizontal string lines were impressed across its face at 1-in. intervals. During this period, there was no perceptible evidence of creep.

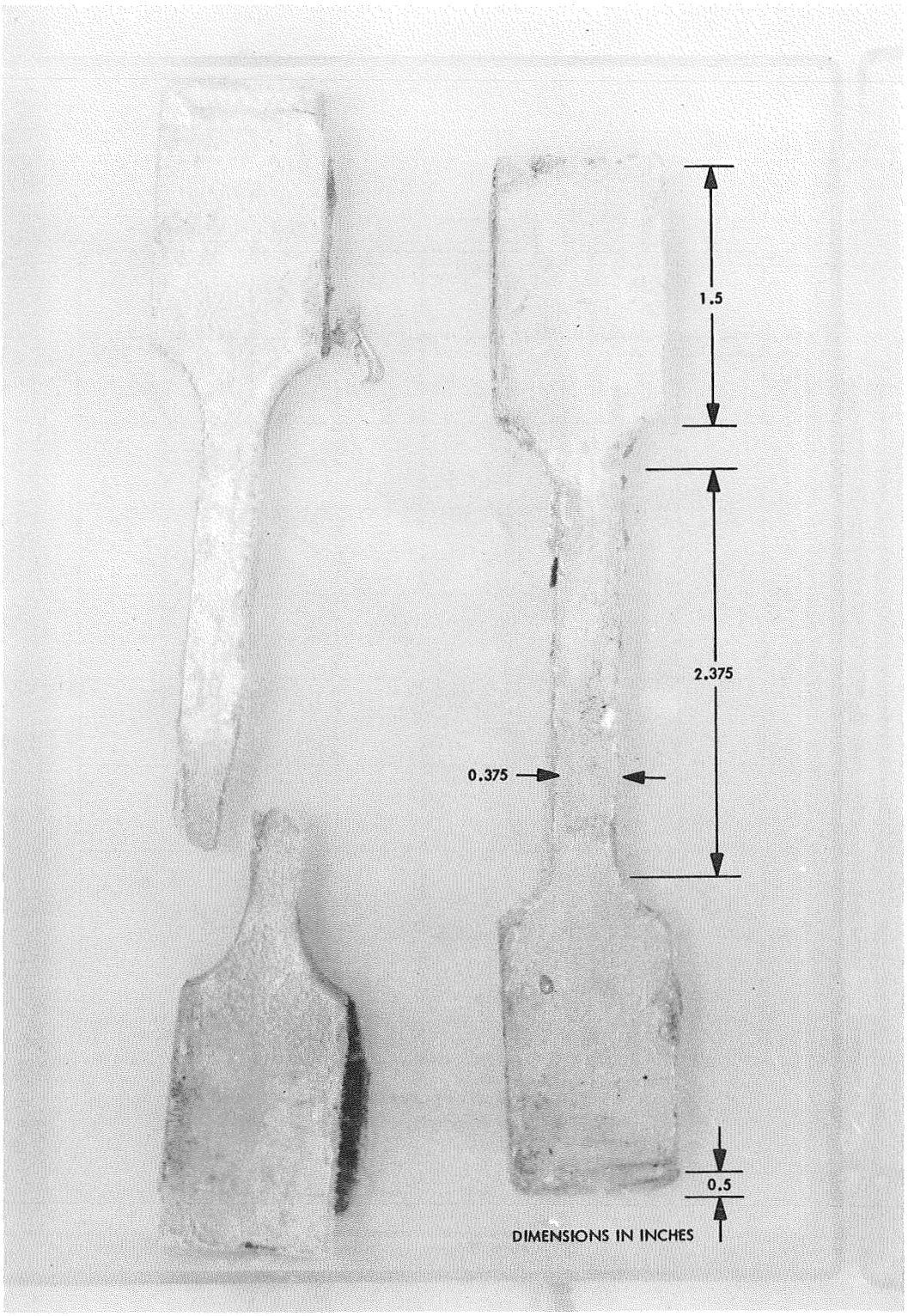


Fig. 5. Comparison of tensile test specimens

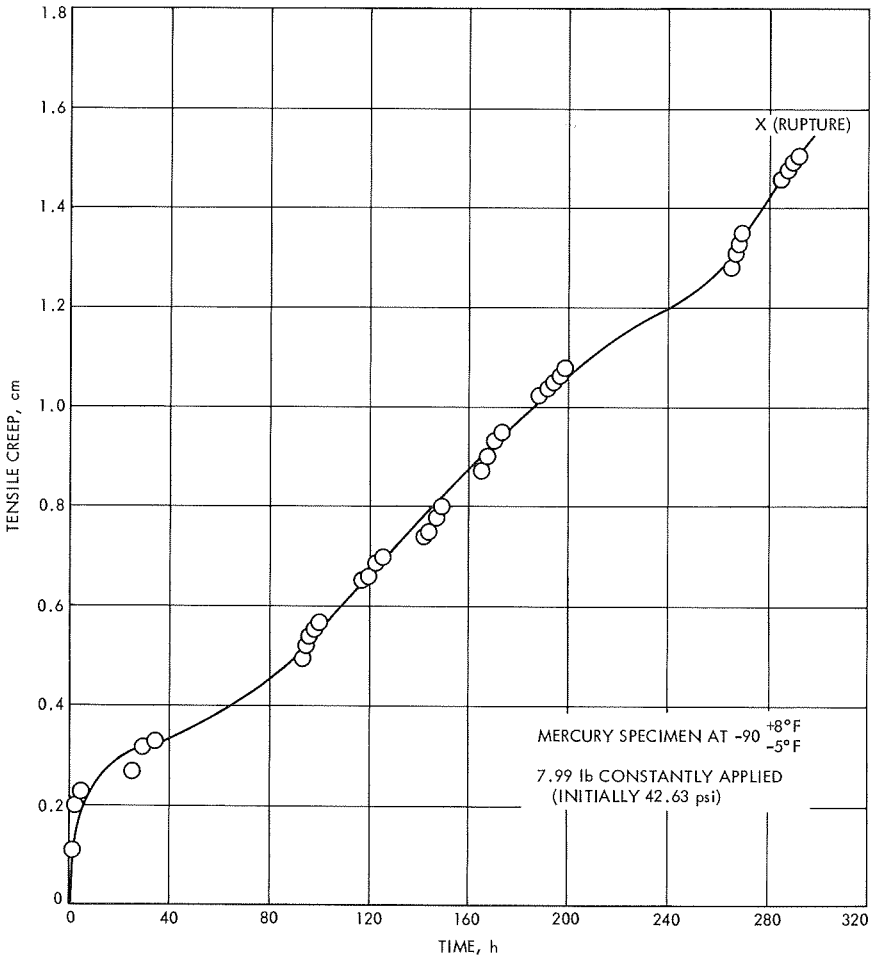


Fig. 6. Tensile creep test

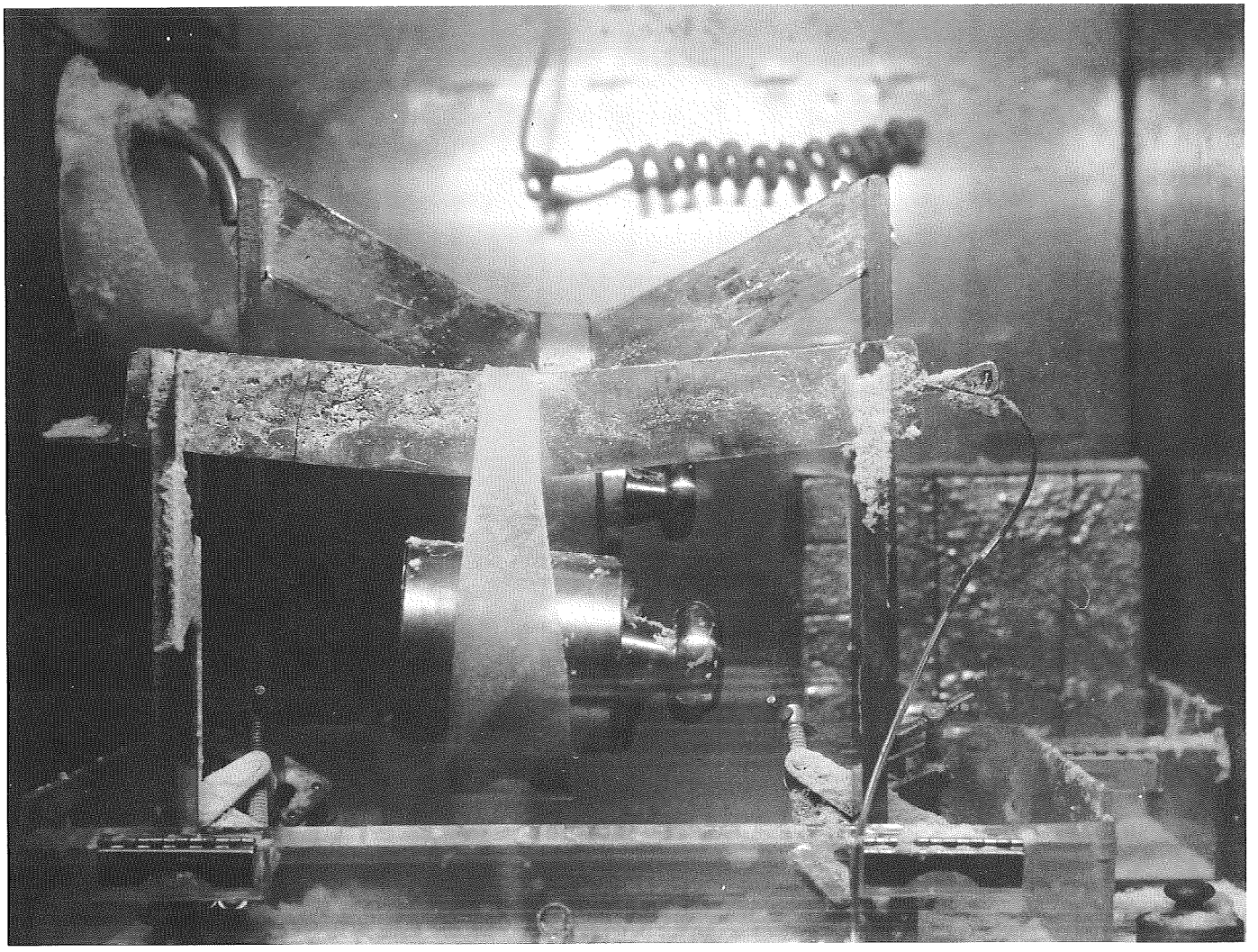


Fig. 7. Beam bending test and static (no externally applied load) test

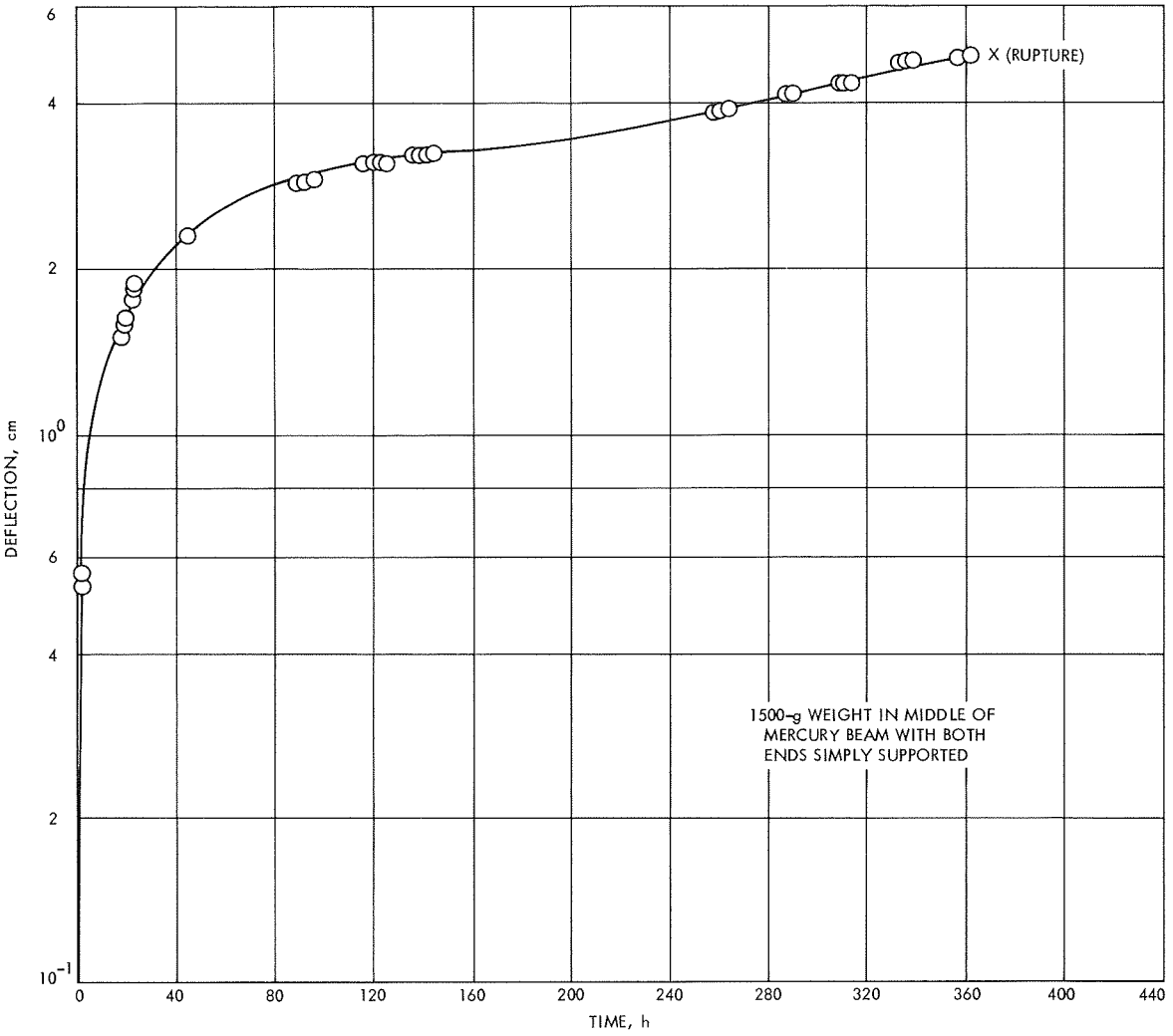


Fig. 8. Deflection creep curve

Several larger mercury slabs have been frozen and kept in the chambers at −95°F. Lines have been scribed at 1-in. intervals and they are being observed for creep.

7. Concluding Remarks

The findings of this program can be summarized as follows:

(1) Liquid mercury adheres poorly to solids, but solid mercury at −95°F adheres resolutely to the container it has been frozen in.

(2) Relatively long-time structural properties of frozen mercury (other than single crystal properties) are not available.

(3) Elasticity of solid mercury was not discernible when using optical instruments that are read to the nearest 0.01 cm.

(4) Creep is relatively large and measurable when compressive loads of 97.4 psi and 37.5 psi are applied to frozen mercury at −95°F.

(5) Application of 42.6 psi tensile stress to a mercury test specimen at −95°F induces creep and specimen rupture.

(6) A 2.5-psi static load produced no measurable creep in frozen mercury at −95°F (within the time limits of this work).

(7) A method for computing induced stresses from frozen mercury beam bending data is available, provided the beam end supports do not induce lateral tensile stresses.

There are grid configurations that minimize the effects of creep. The following tests should be continued to study this design approach:

(1) Varied static load tests for determining grid geometry.

(2) Beam bending tests (with end supports that do not induce lateral tensile stress) so that working stress levels may be computed.

Reference

1. *Standard Handbook for Mechanical Engineers,* Seventh Edition, Section 5-13. Edited by T. Baumeister and L. S. Marks. McGraw-Hill Book Co., Inc., New York, 1967.

XVII. Solid Propellant Engineering

PROPULSION DIVISION

A. High-Speed Streak and Framing Camera Techniques for Electro-Explosive Device Performance Evaluation, *O. K. Heiney, B. Johnson, and J. West*

1. Introduction

The simultaneous streak and framing camera is a useful and versatile tool for examination of the high-rate combustion phenomenon occurring in solid propellants and electro-explosive devices. Proper backlighting of the photographed event will manifest air density variations which will show shock wave formation and travel. From the shock velocity, the pressure output of a squib can be precisely determined and the formation can be studied. The following discussion will describe the optical techniques used and the experimentally determined pressure–distance measurements from a detonator, a squib and an electric match.

2. Technique

The camera has a mirror control pressure system as well as delay and pulse generators for sequencing lighting and event initiation with the camera's rotating mirror. Figure 1 is a schematic of a typical optical setup used for an experimental firing. Back lighting is provided by an exploding wire point light source which is collimated by a condensing lens to provide parallel light at the event. These parallel rays will provide a modified Schlieren effect and render the shock wave visible. Knowing the framing rate or the streak writing speed, the wave-travel velocity is determined and thus the overpressure. The overpressure

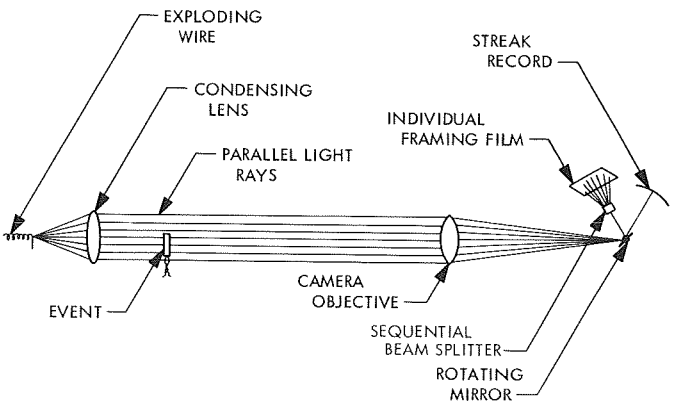


Fig. 1. Typical optical setup for experimental firing

in air as a function of time and distance (Ref. 1) is simply defined as

$$\frac{P_2}{P_1} = \frac{2\gamma M_1^2 - (\gamma - 1)}{\gamma + 1} \tag{1}$$

where

M_1 = Mach No. of wave velocity into undisturbed air

P_1 = ambient undisturbed air pressure

P_2 = shock pressure behind wave front

γ = specific heat ratio for air

Figure 2 illustrates a 12-frame sequence of an *Atlas* match initiation. The 12 frames are sequentially laid out from left to right and from top to bottom. The black line indicates the location of the associated streak data for the event. The streak is fixed in location and varies with time on the film. It is seen that both an initial shock wave and its driving particle velocity wave become visible by the fourth to fifth frame in Fig. 2. Figure 3 is a velocity-distance plot for both of these wavefronts as taken from the framing camera data. The range of initial shock velocities corresponds to monotonically declining delivered pressures from 92 to 40 psia.

Figures 4 and 5 are the framing sequence and associated streak for an initiation of a *Surveyor*-type squib. An interesting feature of this device is the fact that the delivered pressure is constant at 175 psia for the entire field of view, some 1.5 in. Figure 4 also illustrates the

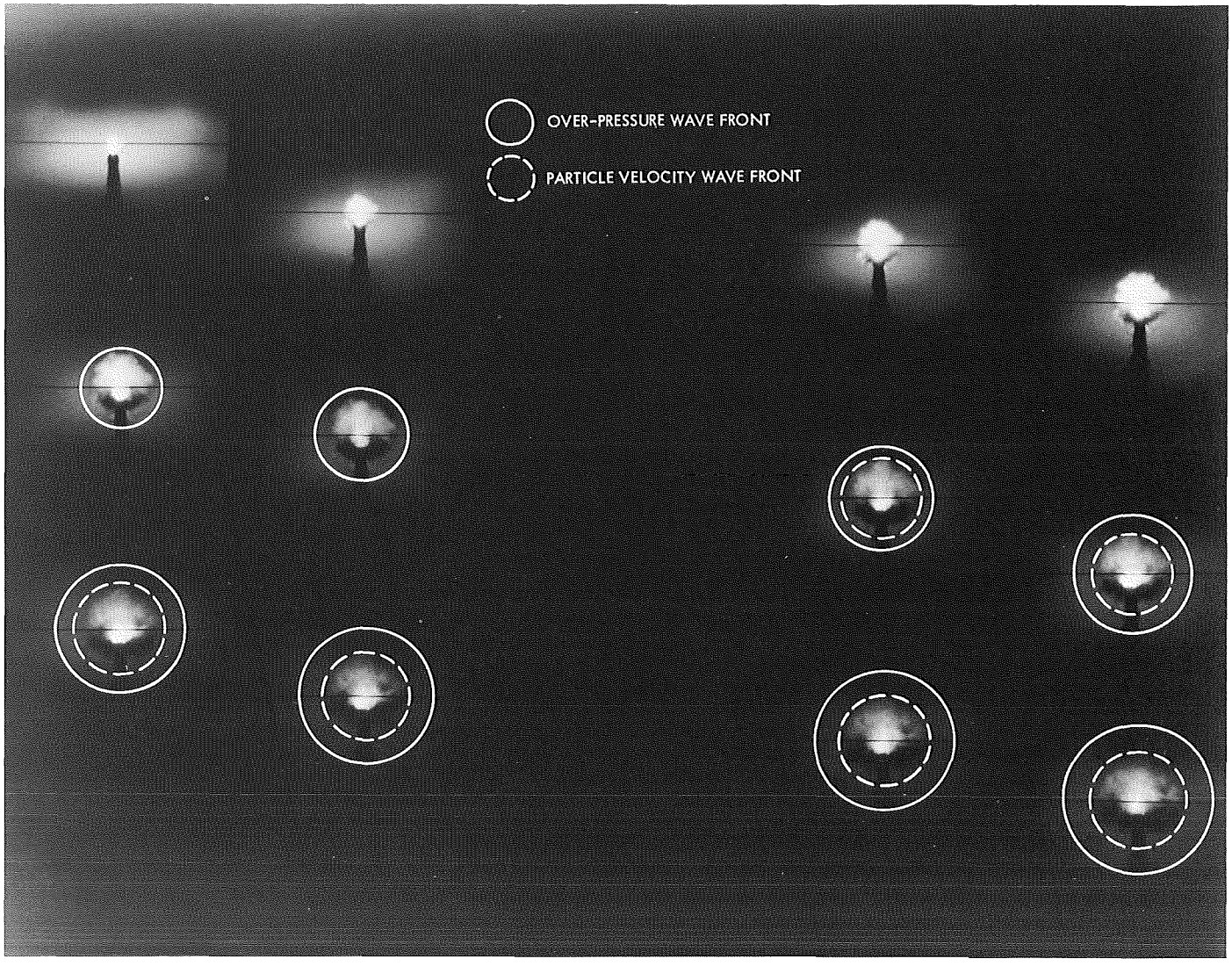


Fig. 2. *Atlas* match initiation 12-frame sequence

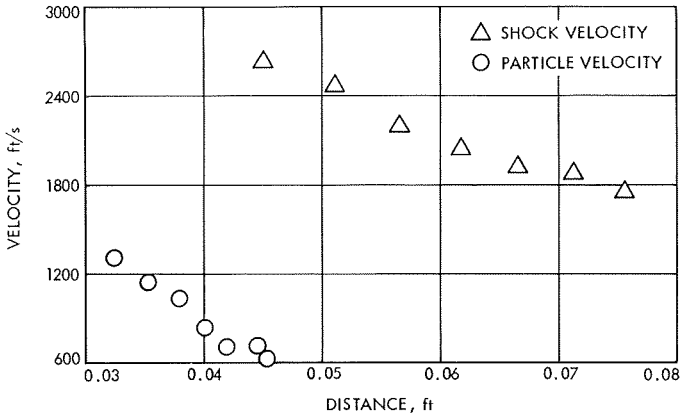


Fig. 3. Velocity versus distance for match initiation

barrel shape of the output from this squib rather than the more typical spherical pattern. The streak of this event (Fig. 5) allows a continuous evaluation of the important velocities but obscures much of what is physically occurring. The framing shots give an excellent insight into the phenomena but allow velocity measurement at only discrete points. Thus the streak and framing pictures are nicely complementary and provide a significant advantage to the simultaneous streak and framing optical setup. For all framing pictures illustrated, the interframe time is 2.5 μs.

A standard dynamite detonator sequence is shown in Fig. 6 with streak and framing data correlation illustrated

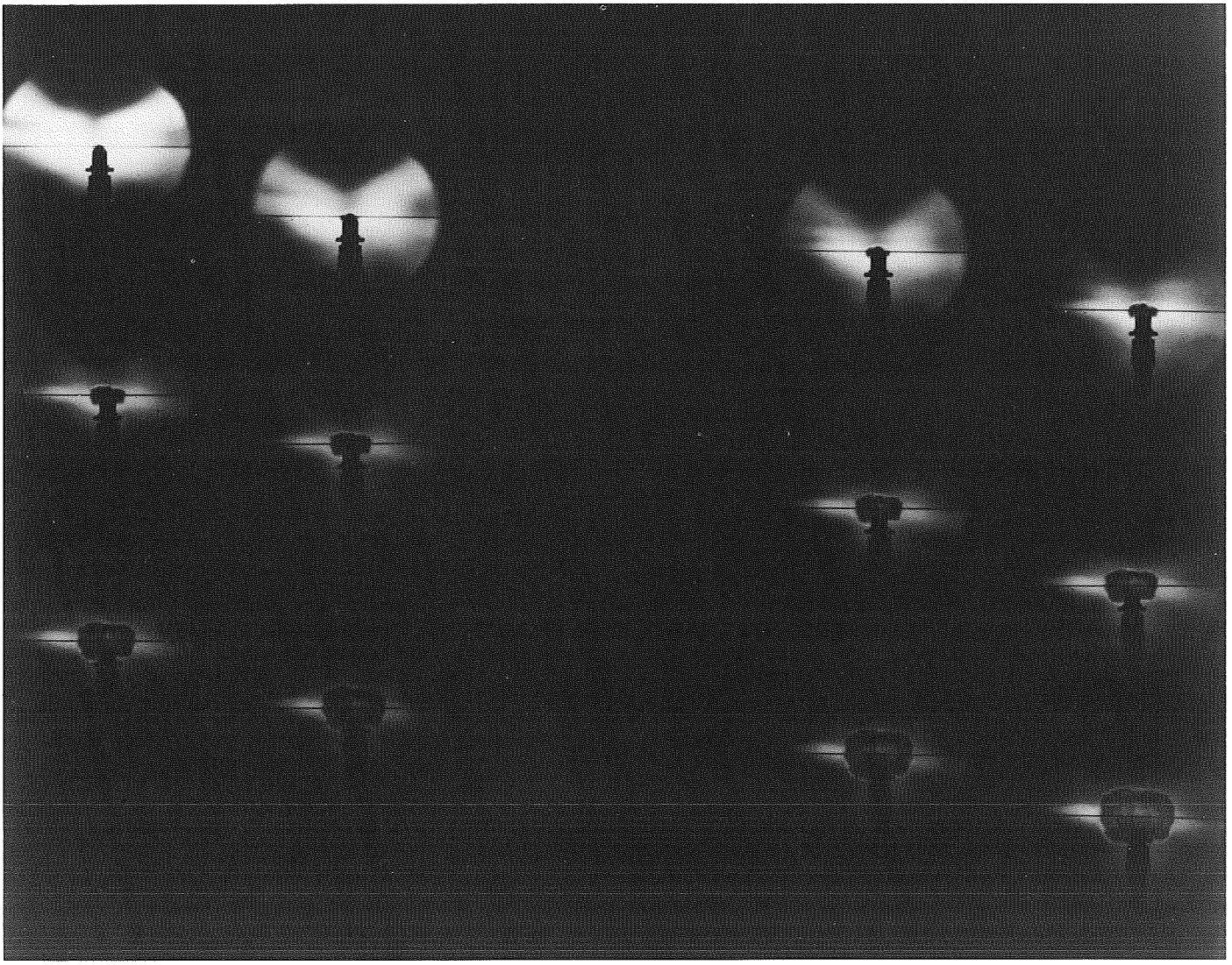


Fig. 4. Framing sequence for squib initiation

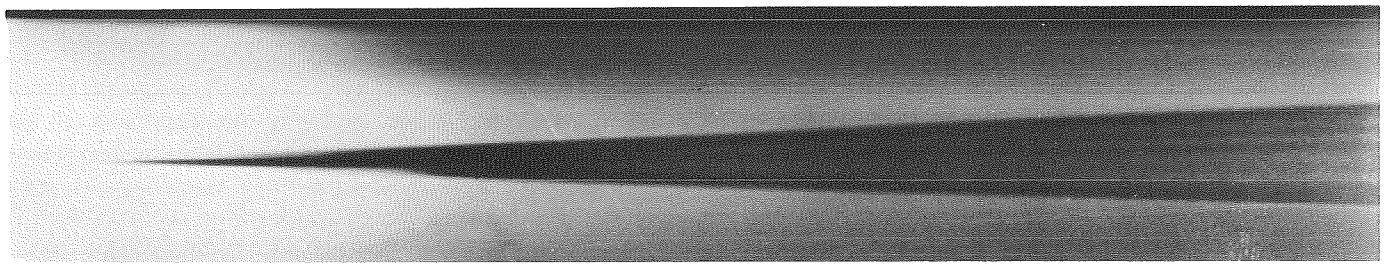


Fig. 5. Streak of squib initiation

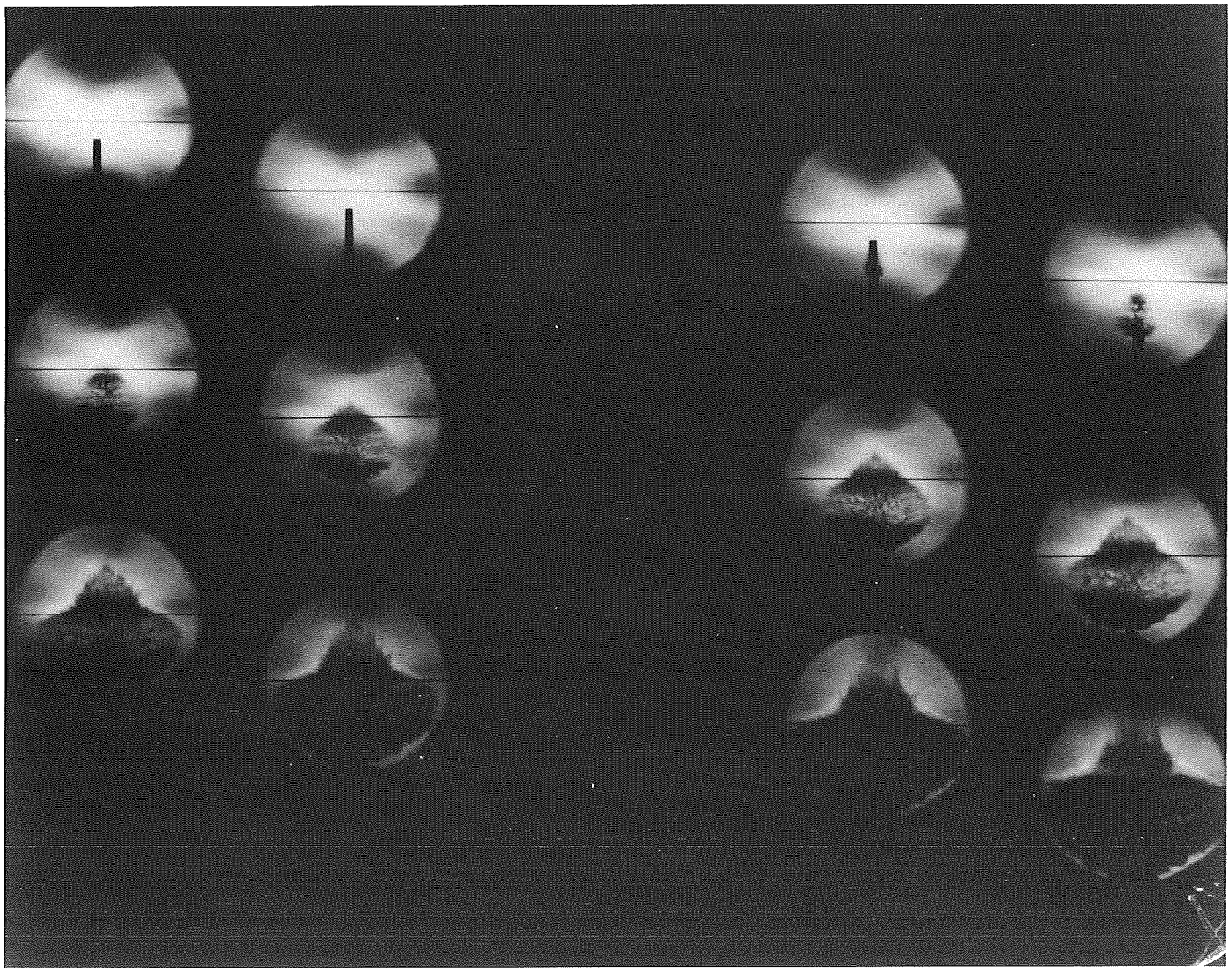


Fig. 6. Standard dynamite detonator sequence

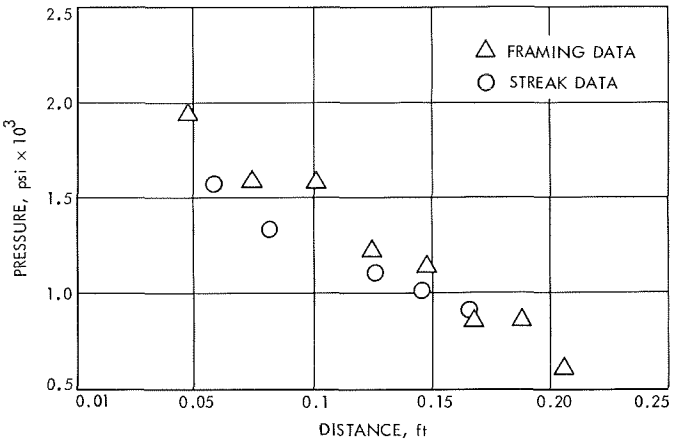


Fig. 7. Pressure versus distance for detonator

by Fig. 7. The detonation phenomenon is occurring more rapidly than the previously discussed deflagrations and the moltent casing is traveling at the blast velocity. The data abstracted from both streak and framing sequence is seen to be in quite good agreement. The blast wave established decays from 2,000 to 600 psia overpressure within 3 in. of travel.

3. Results

In summation, it can be stated that streak and framing camera techniques provide a unique and flexible evaluation tool for electro-explosive device applications, both as a quality-control device and also as a method to gain insight into the physical phenomena occurring during the high-rate reactions involved.

Reference

1. *Equations, Tables and Charts for Compressible Flow*, NASA Report 1135. National Aeronautics and Space Administration, Washington, D.C., 1953.

B. Solid Propellant Rocket Motor Command Termination by Water Injection, *L. D. Strand*

1. Introduction

A study of water quench of solid propellant rockets is being conducted with the objectives of: (1) determining the optimum method of water injection to obtain extinguishment, (2) better understanding the quench mechanism so that it may be optimized, (3) establishing means for predicting water requirements of any given motor, and (4) determining the feasibility and performance of water extinguishment of flight motors. The laboratory motor being utilized, along with its accompanying test apparatus, instrumentation, and test procedure, was described in SPS 37-52, Vol. III, pp. 89–96. The results of extinguishment tests using a sheet injector and overhead atomizer injectors were discussed in SPS 37-55, Vol. III, pp. 169–173.

2. Tests

Since the last report, the window motor quench tests with high-speed motion picture coverage were completed. Nine additional motor firings were carried out, with successful quench occurring in five of the tests. The five successful tests consisted of the following:

(1) Two tests using the overhead injector system with motor pressure and water injection rates of approximately 95 psia and 1 lb/s, respectively. The ratio of water injection rate to the calculated propellant gasification rate was approximately 5.5. Solid-stream and flat-spray water injectors were used.

(2) Two tests using the overhead injector system with motor pressure and water injection rates of approximately 260 psia and 2 lb/s, respectively. The ratio of water injection and calculated propellant gasification rates was approximately 6. Flat-spray and solid-cone water injectors were used.

(3) One sheet injector motor fired at a motor pressure of 150 psia and an injection rate of approximately 2½ lb/s.

The camera field of view and framing rate were approximately 2 in.^{2} and 8,000 frames/s, respectively.

The same nonaluminized polyurethane–ammonium perchlorate propellant formulation was used for all the tests. Plans to include metallized propellants in the window motor test phase were dropped when the first test resulted in extensive condensation of aluminum oxide on the quartz window with some pitting of the window.

3. Test results

The test parameters and reduced results for the five successful quenches (Nos. 20–23, 25), plus those of the earlier quench tests described in SPS 37-52 and 37-55, Vol. III, are listed in Table 1. The results can be summarized as follows:

(1) In general, the injector spray pattern used in the overhead injector system did not produce a readily detectable effect on either the quantity of water injected up to complete quench or the rapidity of the quench, as indicated by the photocell and motor pressure signals.

Table 1. Test parameters and results

Test No.	Nozzle throat diameter, in.	Injector system	Accumulator water pressure, psig	Nominal motor pressure, psia	Water injection rate $\dot{\omega}$, lb/s	$\frac{\dot{\omega}_{water}}{\dot{\omega}_{propellant}}$	Maximum motor dP/dt, psi/s	Water volume injected, ml	$\Delta\tau_1$, ms	$\Delta\tau_2$, ms	ω_1, ml	ω_2, ml	Test results
9	0.600	Flat spray 8 elements	300	95	4.8	25	6,000	84	—	—	7	—	Quench
12	0.640	Fog 8 elements	1900	75	14.1	82	7,600	90	4.2	10.0	28	51	Quench
14	0.620	Sheet	300	81	13.7	77	12,000	175	2.4	5.0	14	28	Quench
15	0.620	Sheet	310	77	3.5	21	5,200	70	12	6.7	17.5	24.5	Quench
16	0.620	Fog 8 elements	285	72.5	6.0	36	1,900	56	16.5	19	42	56	Quench
17	0.480	Fog 8 elements	390	180	8.3	33	1,900	63	13	—	49	—	No quench
20	0.600	Solid stream 8 elements	120	92	1.0	5.5	1,200	56	24	58	7	45.5	Slow quench
21	0.600	Flat spray 8 elements	125	98	1.1	5.5	690	50	25	94	7	49	Very slow quench
22	0.440	Flat spray 8 elements	290	≃265	1.8	≃6.0	—	98	24	41	14	50	Quench
23	0.440	Sheet	400	150	8 to 2.4	—	10,000	63	14	29	28	63	Marginal quench
25	0.440	Solid spray 8 elements	280	260	1.8	6.4	3,600	70	25	38	14	53	Quench

$\Delta\tau_1$ = time between initiation of water pressurization and initiation of photocell signal drop.
$\Delta\tau_2$ = time between initiaton of photocell signal drop and leveling off of signal.
ω_1 = quantity of water injected up to initiation of photocell signal drop.
ω_2 = quantity of water injected up to leveling off of photocell signal.

(2) At a constant chamber pressure, increasing the rate of water injection decreased the quantity of water required to produce a quench and increased the rapidity of the quench.

(3) With approximately equal injection to propellant gasification rate ratios, roughly equal quantities of water were injected up to complete quench at the two motor chamber pressures tested, i.e., the ratio correlated the data for these few tests.

(4) The motion picture data showed quenching of combustion (darkening of combustion zone above propellant surface) to be coincident with the advancing of the water interface over the propellant surface. As stated in SPS 37-55, Vol. III, pp. 169–173, simple thermal quenching of the propellant combustion zone and continued cooling of the surface until the hot gases have been exhausted seems to be a reasonable model for describing the extinguishment characteristics.

4. Test System Modifications

During this test program it became obvious that one of the program objectives—a determination of the optimum method of water injection to obtain extinguishment—could not be determined from these tests. It was decided that, in order to determine any possible differences between the different injector systems, a series of *go, no-go* type tests were required using more carefully controlled injected water quantities and injection rates than were possible with the existing test system. The test plan has been amended to include such a series.

The test system has been modified and is in the process of being checked out. The ac solenoid valve has been replaced with two dc Annin valves, one normally closed and one normally open. The valves are attached to a 5-gal surge tank, pressurized to 150 psi. Actuation of the valves is controlled by two digital counters. To reduce the water pressure drop during injection, the 5-gal accumulator has been reinstalled. The quantity of water injected is controlled by the water pressure and the time interval between actuation of the valves. The results of initial checkout tests have been encouraging.

5. Future Tests

For each injector system the added tests will consist of a systematic reduction of the weight of water injected

at a given chamber pressure and water injection rate to determine the minimum quantity of water required to produce a quench. The latter two parameters will be varied over a limited region. Following completion of these tests, equipment buildup for the Phase 2 motor tests will be initiated.

C. T-Burner Investigation, *E. H. Perry*

1. *Introduction*

In a previous article (SPS 37-54, Vol. III, pp. 93–96) the T-burner and its role in combustion instability studies were discussed along with some results obtained in T-burners at JPL. Further tests have now been completed and have provided additional information on the nature of the burner and its value in combustion instability investigations.

2. Transparent T-Burner Studies

A series of tests using transparent plastic chambers was carried out to permit high-speed motion photography of T-burner firings. Lucite chambers with an inner diameter and wall thickness of 1.5 and 0.25 in., respectively, and lengths varying from 7.0 to 13.0 in. were used. The frequency of the pressure oscillations, determined by the chamber length, ranged from approximately 1000 to 2200 Hz. The camera framing rate of 8000 frames/s allowed a minimum of four frames during each cycle of the oscillation. A composite polysulphide propellant denoted as T-17 was used predominately, since it provided pressure oscillations of moderate amplitude with little tendency toward forming dark deposits on the chamber walls. Figure 8 shows a transparent burner mounted on the surge tank used to provide a constant chamber pressure of 100 psig.

These tests demonstrated rather clearly that the T-burner is, indeed, a one-dimensional device with three-dimensional effects confined primarily to the vicinity of the vent. The combustion gases issuing from the burning surface appeared to flow rather smoothly toward the vent with large-scale circulation effects absent. The temperature fluctuations accompanying the pressure oscillations were readily observed as pulsations of light at the ends of the chamber. Like the pressure oscillations, the temperature fluctuations at the two ends were approximately 180 deg out of phase with each other.

In addition to the observations of the flow field and temperature fluctuations one of the more interesting results of this series was a comparison of two ignition techniques. One technique resulted in rather poor ignition, whereas the other provided a rapid even ignition of the entire propellant face. Since the T-burner method depends to a large extent on an even ignition and burnout of the propellant, such a comparison seemed highly relevant.

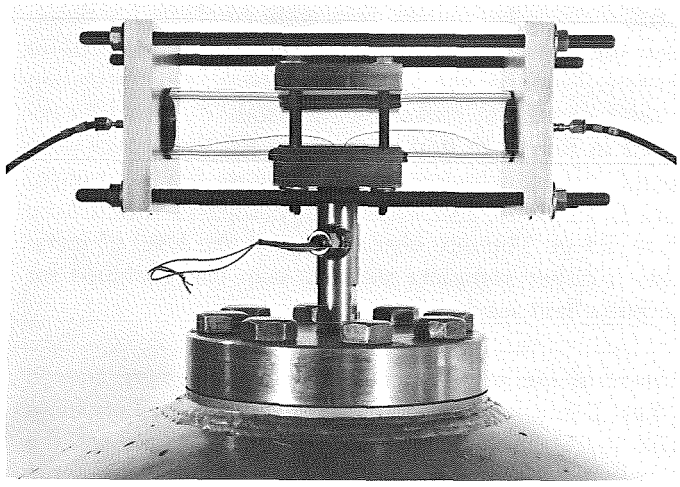


Fig. 8. Transparent T-burner mounted on surge tank

In both techniques compared, a pyrotechnic ignition paste was applied to the surface of the propellant, since previous studies had indicated that rather poor ignition resulted otherwise. In the first technique a Nichrome igniter wire was simply bonded to the pyrotechnic paste, whereas in the second method the Nichrome was imbedded within a small tablet or teardrop of the same paste. This was, in turn, bonded to the propellant.

The motion pictures showed that in the first case, ignition occurred very rapidly near the igniter wire but rather slowly over the remainder of the propellant face. This is in contrast to the second method, which provided a rapid ignition of the entire surface. Apparently the difference between the two lay in the fact that ignition occurred not as a result of flame spread across the ignition paste but rather as a result of sparks arising from the burning paste striking unburned regions. In the first technique sparks were simply spewed out away from the propellant surface, while in the second method many were directed back toward the surface from the burning igniter tablet. Tests using this latter technique generally resulted in better reproducibility of practically all observable data—another indication of even ignition, and hence even burnout, of the propellant disks.

3. Application of T-Burner Results

Since the primary function of the T-burner is to provide a laboratory tool for investigating combustion instability

in full-scale rocket motors, tests were conducted to determine the extent to which this connection is realized in practice. Probably the most comprehensive investigation of solid propellant rocket motor instability was that reported in Ref. 1. In that study some 400 test firings were carried out in cylindrical motors whose lengths ranged from 17 to 44 in., and whose port diameters varied from 1.5 to 5.5 in. Through observations of the level of instability a rather complete stability map was obtained for the propellant used.

Some of the results of this investigation were treated in Ref. 2 in a general analysis of acoustic oscillations in solid rocket motors. In this treatment, values for the propellant admittance function were deduced from the stability boundary of Ref. 1 and an assumption regarding the principal acoustic loss mechanism. The propellant admittance function, or the related response function, describe the manner in which the burning propellant surface responds to a pressure perturbation. The T-burner hopefully measures this response function and thereby provides data relevant to instability studies in full-scale rocket motors.

Recent tests at JPL measured for the first time the response function of the propellant used in Ref. 1. The tests were conducted in 1.5-in.-dia T-burners at mean chamber pressures of 300, 400, and 600 psig and an initial grain temperature of 160°F. The growth-decay T-burner method was used, since the propellant did not oscillate well enough to use the variable-area method. In Fig. 9 the data obtained in these tests are denoted by the various symbols and plotted as a function of oscillation frequency. The response function is a complex quantity, and it is only the real portion that is shown here. It is the real part that determines the stability characteristics of the propellant and, hence, is of primary interest here.

In addition, Fig. 9 also shows the values of the response function deduced in Ref. 2. These are indicated by the dashed line. Actually in Ref. 2 the admittance function A_b was calculated, but it is related to the response function R_b by the following equation:

$$R_b = \frac{A_b + 1}{\gamma} \tag{1}$$

where γ is the usual specific heats ratio of the combustion gases.

While the agreement between the predicted curve and the experimental data is not excellent, it is most encouraging and indicates that the T-burner may, indeed, provide data that is very pertinent to the problem of combustion instability in actual rocket motors. The moderate agreement observed only underscores the need for more comparisons of this type, for it is only through such comparisons that a true assessment of the value of the T-burner in such studies can be realized.

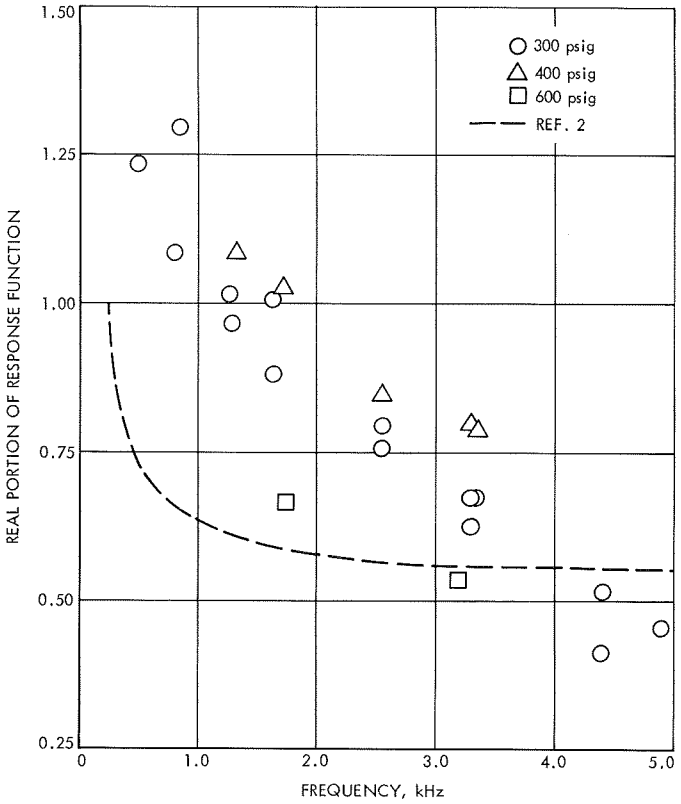


Fig. 9. Predicted and observed values for the response function of the propellant used in Ref. 1

4. Studies of the Variable-Area Method

In the SPS article previously mentioned, two methods of conducting T-burner tests were discussed and data obtained by the two methods was compared. It was found that the variable-area method consistently indicated lower values for the propellant response function than did the growth-decay method. In an attempt to explain this lack of agreement, additional tests were conducted using the variable-area method. They indicated that data obtained through this method may be questionable.

The tests were carried out in 1.5-in.-dia T-burners at a chamber pressure of 300 psig and a grain temperature of 80°F. The propellant, obtained from the Naval Weapons Center at China Lake, Calif., was a composite denoted

as A-13. In variable-area tests the area ratio referred to is simply the ratio of the area of the propellant disk to the cross-sectional area of the chamber. For the present tests area ratios of 0.43, 0.72, and 1.00 were used. Naturally, an area ratio of unity indicates that the propellant disk has the same diameter as the chamber. Area ratios less than unity were obtained by bonding disks of propellant into stainless steel rings which were, in turn, bonded to the chamber walls.

In the variable-area method the oscillation growth rate is measured as a function of area ratio. If the resulting growth constants are plotted against the area ratio, a straight line should pass through the data. The slope of this straight line provides a measure of the propellant response function as described in Ref. 3.

In the present tests, data were obtained for oscillation frequencies of 850, 1700, and 3400 Hz. At least two tests at each of the three area ratios were carried out for each frequency, to provide an indication of the reproducibility of the data. In general, the reproducibility was rather good, as is shown in Fig. 10 where the growth constants for the tests conducted at a frequency of 800 Hz are plotted against the area ratio. It is apparent that the data do not lie along a straight line and, hence, do not provide a unique value for the response function. In fact, if straight lines are drawn first through the data at high area ratios (0.72 and 1.00) and then through that at low ratios (0.43 and 0.72) and the corresponding slopes are computed, response values of 1.87 and 1.04, respectively, result. This is also indicated in Fig. 10. The data obtained at the two other frequencies exhibit a behavior very similar to that displayed in Fig. 10. In all three cases the data do not lie along a straight line, which raises rather serious doubts about the usefulness of this method. Although the reason for this behavior is not fully understood at the present, it is felt that it is due, in part, to the fact that the losses in the chamber may depend on the area ratio. The data appear to confirm this possibility. Regardless

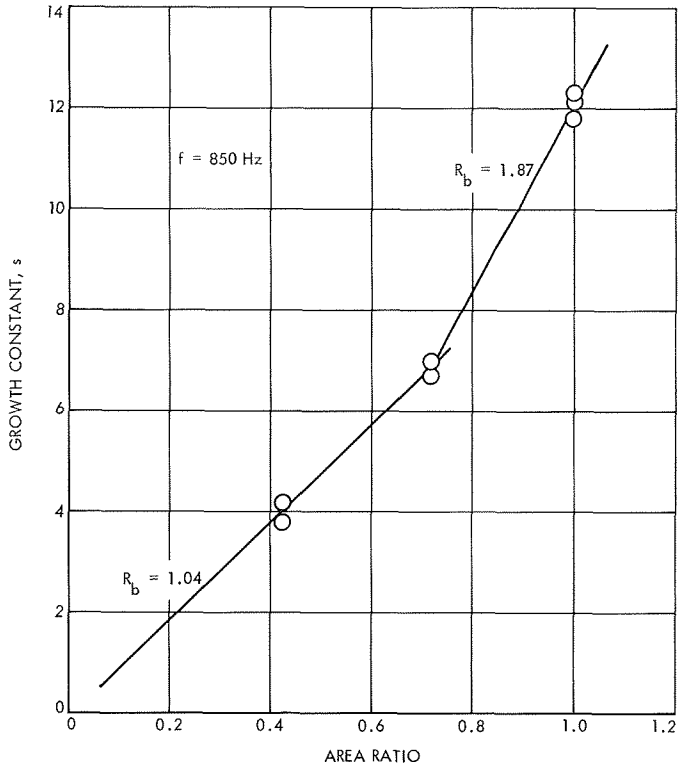


Fig. 10. Growth constants obtained by the variable-area method plotted against area ratio

of the cause behind the behavior, the fact remains that the variable-area method appears to provide rather questionable data.

References

1. Brownlee, W. G., and Marble, F. E., "An Experimental Investigation of Unstable Combustion in Solid Propellant Rocket Motors," *Prog. Astronaut. Rocket.*, Vol. 1, pp. 455–494.
2. Culick, F. E. C., *"Acoustic Oscillations in Solid Propellant Rocket Chambers," Astronaut. Acta*, Vol. 12, pp. 113–126, 1966.
3. Beckstead, M. W., and Culick, F. E. C., *A Comparison of Analysis and Experiment for Solid Propellant Combustion Instability*, Technical Publication 4531, p. 13, Naval Weapons Center, China Lake, Calif., May 1968.

XVIII. Polymer Research

PROPULSION DIVISION

A. Behavior of Polymeric Materials in Various Environments, *S. H. Kalfayan and R. H. Silver*

1. Introduction

In a previous report (SPS 37-57, Vol. III, pp. 160–163), the effects of a simulated Venus atmosphere on some selected polymeric materials were discussed. This atmosphere consisted mostly of carbon dioxide, i.e., 98.6%. Other constituents were: water, 0.1%, oxygen, 0.8%, and nitrogen, ca 0.5%. The temperature of the simulated atmosphere was held at 550 $\pm$10°F and the pressure at 18.5 $\pm$1 atm.

Of the three factors, temperature, pressure, and chemical environment, pressure would be expected to have a negligible effect on the mechanical and physical properties of the test specimens. The other two factors are likely to change some of these properties.

The effects of air, inert atmosphere, such as nitrogen, and vacuum on polymeric products at elevated temperatures have been widely studied. It was of interest, therefore, to compare the effects of the "Venus" atmosphere, composed essentially of carbon dioxide, with those of air and nitrogen at 550°F and ca 18 atm. The results of experiments carried out to make such a comparative study are discussed below.

2. Experimental Method

Three commercially available materials of known heat resistance were selected for evaluation: Kapton 100, a polyimide film manufactured by DuPont; Fiberglas 91 LD, a glass-reinforced phenolic resin made by American Reinforced Plastics Co., and Exp 820, a glass-cloth-reinforced benzimidazole, manufactured by Whittaker Corp.

Exposures to air, nitrogen, and carbon dioxide ("Venus") were carried out in a special pressure chamber described previously (SPS 37-57, Vol. III, pp. 157–160). The materials were tested according to standard methods, with the exception of the tensile test for Exp 820. Because of insufficient material, tensile bars made of this product were not of standard size. For the "Venus" and nitrogen exposures, the test chamber was evacuated and flushed several times with the test gas, before being filled. A

slow flow of gas was maintained through the chamber, during the course of the exposure. The experimental conditions were:

Pressure, atm	18.5 ±1.0
Temperature, °F	550 ±10
Exposure period (at above conditions), h	24

3. Results and Discussion

Some values obtained after the various exposures are given in Tables 1 and 2. For the sake of brevity, results are reported, in most cases, as percent of retained properties or percent change. Values shown are averages of at least three tests. In reporting percent of retained properties, numbers larger than 100 indicate that the particular property was enhanced after exposure. Comparisons were made to unexposed, control samples.

Results shown in Table 1 indicate that the "Venus" environment with the carbon dioxide atmosphere, containing some water vapor and some oxygen, reduced the mechanical properties of the three materials tested, more than air or nitrogen. It should be noticed, however, that, with the exception of percent elongation, the mechanical properties were not reduced by more than 10%. The improvement in the tensile strength of Kapton and Fiberglas 91 LD in nitrogen is worthy of notice as is the tensile property of Exp 820 in air.

No dramatic changes in the volume resistivity of Kapton and Fiberglas 91 LD were encountered after any of the exposures. Shortage of Exp 820 prevented the performance of the volume resistivity measurement. The weight of Kapton increased after the "Venus" exposure. Also, the weight loss of the other two materials was less after exposure to this atmosphere. Absorption or an actual chemical combination between the imide nitrogen

Table 1. Retained mechanical properties after 24-h exposure to various gaseous environments

Material	Tensile strength, % retained			Elongation, % retained			Hardness unit change (Rockwell E)			Tear strength, % retained		
	Air	N_2	"Venus"	Air	N_2	"Venus"	Air	N_2	"Venus"	Air	N_2	"Venus"
Kapton 100 (polyimide film)	99	111	90.6	78.6	127	54	—	—	—	92.7	94.6	89.5
Fiberglas 91 LD (glass-filled phenolic)	108	122	94	107	113	77	+1.5	+1.5	+2.0	—	—	—
Exp 820 (glass-fiber-reinforced benzimidazol)	143	115	95	65	60	60	+7	+7	+4	—	—	—

Table 2. Change in properties after 24-h exposure to various gaseous environments

Material	Volume resistivity, Ω-cm				Weight change, %			Volume change, %		
	Control	Air	N_2	"Venus"	Air	N_2	"Venus"	Air	N_2	"Venus"
Kapton 100 (polyimide film)	1.24×10^{17}	5.6×10^{16}	8×10^{16}	1.20×10^{17}	−0.1428	−1.373	+1.950	−0.98	−1.36	—
Fiberglas 91 LD (glass-filled phenolic)	1.44×10^{14}	1.0×10^{14}	9×10^{14}	1.03×10^{14}	−1.859	−2.389	−0.1488	+1.49	−0.28	0.00
Exp 820 (glass-fiber-reinforced benzimidazole)	—	—	—	—	−1.671	−1.541	−1.202	0.00	−1.64	−2.52

of the Kapton and the acidic carbon dioxide is a possibility.

Although the "Venus" atmosphere affected the properties of the tested materials more than air or nitrogen did, its effect was not severe enough to rate these three materials to be incompatible with this atmosphere.

B. Long-Term Aging of Elastomers: Chemical Stress Relaxation Studies of Peroxide-Cured Styrene–Butadiene Rubber, *S. H. Kalfayan, R. Rakutis, and R. H. Silver*

1. Introduction

Predicting the long-term behavior of materials, particularly that of elastomers, is of paramount importance. A comprehensive theory has been developed (Refs. 1 and 2) whereby the mechanical properties of solid composite propellants and their elastomeric binders can be predicted from typical molecular parameters. In its present form, however, this theory does not account for chemical reactions which are too slow to be determined from short-time measurements. Ascertaining the nature and extent of the chemical changes that take place in elastomers at long times is the general objective.

At the present time, there are few methods for obtaining data of fundamental significance concerning the chemical changes (aging) taking place in elastomers. The most valuable technique, so far, to obtain such data is that of stress relaxation. Developed by Tobolsky and others (Ref. 3), this method yields information about network degradation in rubbers.

The method was, therefore, selected to study aging of rubbers, in the present instance, the aging of styrene–butadiene rubber (SBR).

It has been reported by Thirion (Ref. 4) that relaxation of natural rubber vulcanizates is due to a combination of two factors: aging and viscoelasticity. For natural rubber cross-linked or vulcanized with dicumyl peroxide (DiCup), it was concluded that aging prevails above and viscoelasticity below 65°C. A linear relationship of the tensile force f with time t was observed according to the following relaxation law (Ref. 4):

$$\frac{f}{f_0} = 1 - \frac{t}{a} \tag{1}$$

This relationship was derived from an earlier theory of chain scission in a molecular network (Ref. 5):

$$\frac{f}{f_0} - f = \frac{a}{t} - b \tag{2}$$

where f is the tensile force at time t, and f_0, a, and b are constants. In certain cases, this relationship appears to explain the kinetics of stress relaxation more satisfactorily than an exponential function previously proposed by Tobolsky (Ref. 6).

Stress relaxation rates agreed with Eq. (1) over a wide range of time when temperature was above 65°C, i.e., with aging as the predominant cause of relaxation (Ref. 4). The energy of activation ΔH for this process was calculated from the classical Arrhenius equation:

$$a = a_0 e^{\Delta H/RT} \tag{3}$$

where T is the absolute temperature, R is the gas constant, a is the time constant, and the reciprocal of relaxation rate values, a_0 and ΔH are constants related to the aging of the rubber vulcanizate.

In the present work, similar investigations are being carried out using SBR cross-linked with DiCup. The kinetic analysis of stress relaxation is of primary interest in this study, with the purpose of understanding the basic mechanism of the aging of SBR.

2. Experimental Method

a. Materials. The SBR (lot 386, type 1500) used in this study was obtained from the National Bureau of Standards. It had 23.6% bound styrene, 1.45% antioxidant, tris-(nonyl phenyl) phospite (Poly-Gard), and 6.46% fatty acid soap as emulsifier. The DiCup was obtained from Hercules Powder Co.

b. Compounding and vulcanization. The SBR and DiCup were mixed on a laboratory mill. Batches 2 and 5, discussed here, contained 0.3 and 0.69% DiCup, respectively. The mixtures were molded into 0.010–0.014-in. thick sheets in a press. Vulcanization was effected by heating the sheets at 310°F for 1 h under high pressure.

c. Stress relaxation measurements.

Apparatus. The relaxometer used for the measurement of stress decay shown in Fig. 1, consisted of a universal transducing cell (Statham Model UC3) adapted to 0.5 lb

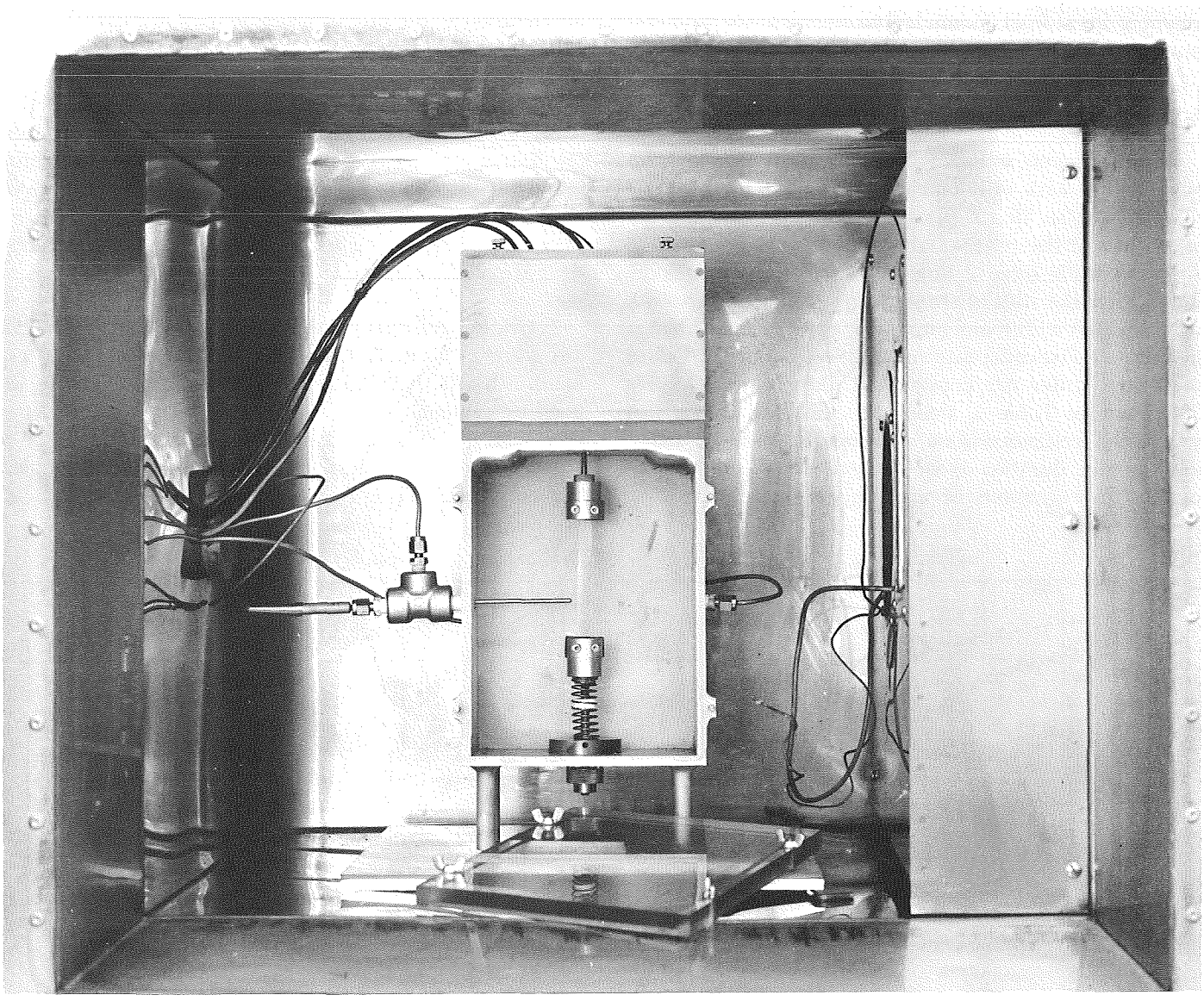


Fig. 1. Stress relaxometer

full-scale tension operation by means of a load cell accessory (Statham Model UL4). The nonlinearity and hysteresis of the system was less than ±0.25% full scale, and was operable from −65 to 250°F. To minimize the thermal drift of the cell, the unit was further protected by surrounding it with a water-cooled coil. Thus, the loadcell temperature would be kept below 90°F while that of the test sample could be at 300°F. The sample and the grips were housed in a 4 × 6 × 8-in. gas-tight chamber, one wall of which was constructed of transparent polycarbonate for observation. The relaxometer was then placed in an air-circulated, temperature-controlled oven. A recorder attached to the relaxometer loadcell gave direct readings of tensile force as a function of time.

Conditions. The SBR test specimens were rectangular pieces, 2.9 × 0.375 in., and were secured to the relaxometer by rectangular grips at right angles to the sample. The specimens were heated in air at a temperature range of 43–115°C, at 29% extension, until complete relaxation took place. Measurements were also carried out in nitrogen at 43, 100, and 115°C. For the latter experiments, the relaxometer chamber was evacuated and flushed several times with high-purity nitrogen. A steady flow of the inert gas was then held during the entire course of a run. Mass spectrographic analysis of the effluent gas from the chamber did not indicate the presence of oxygen either at the start or at the end of each run in nitrogen. In all cases, test specimens were kept at rest in the

chamber for at least 20 min before stressing, for temperature equilibration.

3. Results and Discussion

a. Experiments in air. The results of stress relaxation experiments with SBR batch 2 and batch 5 at various temperatures in air are shown in Fig. 2. Plotted are the logarithms of the relative relaxation f_t/f_0 vs time, where f_t is the tensile force at time t and f_0 is the tensile force 3 min after the start of stressing. After an initial period, which varied with temperature, the rates of stress relaxation fitted a first-order rate equation satisfactorily. Values of the first-order rate constants at various temperatures are shown in Table 3, along with the time intervals at which the relaxation data indicate first-order kinetics, and the total duration of each experiment. After approximately three quarters of each run was completed, an acceleration in the rate of relaxation and a change in the reaction order was observed.

Energy of activation, ΔH values were obtained from Arrhenius plots by using three different sets of data: (1) first-order rate constants; (2) half-times; (3) shift factors, log k_t, along the time axis. The shift factors used are listed in Table 4. They were obtained by superposition of f_t/f_0 vs log t curves at various temperatures over the curve obtained at 115°C.

Activation energies obtained by these three methods are given in Table 5. Except for the value obtained for batch 5 by the half-life method, the other values are in good agreement.

When the shift factors in Table 4 are plotted against the corresponding $1/T$, the graph shown in Fig. 3 is

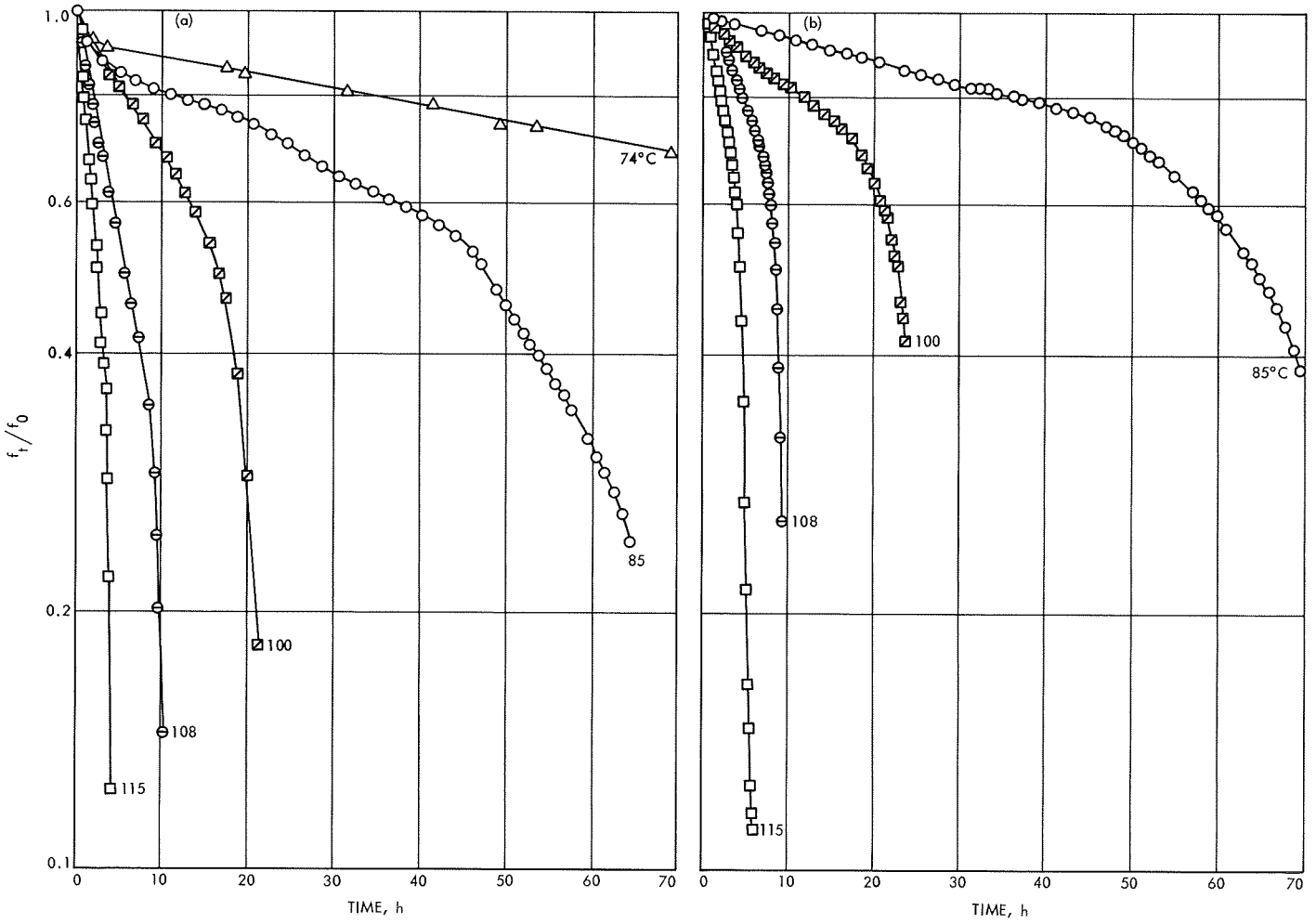


Fig. 2. Effect of temperature on stress relaxation of peroxide-cured SBR: (a) batch 2, (b) batch 5

Table 3. First-order rate constants for the oxidative aging of SBR

Batch	Temperature, °C	K, h^{-1}	σ	Period considered	Total duration, h
2	74	5.40×10^{-3}	0.0002	30–101 h	110
	85	1.45×10^{-2}	0.001	12–48 h	65
	100	3.89×10^{-2}	0.0008	6–16 h	22
	108	1.30×10^{-1}	0.005	3–9 h	10
	115	2.93×10^{-1}	0.003	0.8–3 h	4.2
5	85	6.43×10^{-3}	0.0002	12–48 h	70
	100	1.93×10^{-2}	0.0003	10–17 h	24
	108	1.0×10^{-1}	0.006	4–7 h	9.5
	115	2.32×10^{-1}	0.02	2–4 h	5

obtained. Extrapolation to 27°C (300°K) gives $\log k_t = -5$. According to Thirion,[1] relaxation caused by aging at 27°C and at time t will be the same at 115°C at the end of time t', or

$$\log t' = \log t - 5 \tag{4}$$

$$t' = \frac{t}{10^5} \tag{5}$$

For example, after 100 h at 27°C, the effect of aging on stress relaxation will be equal to that at 115°C after

$$t = \frac{10^2}{10^5} = 10^{-3} \text{ h or } 3.6 \text{ s}$$

It has been pointed out that it is impossible to determine relaxation due to chemical changes (aging) for such a short period. Moreover, the amount of aging during such a short time may be negligible. It follows then, that for the first 100 h or so, stress relaxation of SBR at 27°C may be attributed to viscoelasticity.

[1]Thirion, P., unpublished results.

Table 4. Shift factors along the horizontal logarithmic time axis

Batch	Temperature, °C	Shift factor, log K_t
2	85	−1.3222
	100	−0.9294
	108	−0.4771
	115	0
5	85	−1.6021
	100	−1.1303
	108	−0.5315
	115	0

Table 5. Activation energies, ΔH from different methods for peroxide-cured SBR

Batch	ΔH from first-order kinetics, cal	ΔH from half-life method, cal	ΔH from shift factors, log K_t, cal
2	27,500	27,500	27,900
5	32,800	28,500	33,800

Assuming that for the first 6 s at 115°C aging is negligible, times can be calculated for other temperatures when aging is negligible and stress relaxation is due to viscoelastic behavior only. Table 6 gives the results of such calculations.

It was observed that benzene-extracted samples of DiCup-cured SBR relaxed more quickly than unextracted samples. This is shown in Fig. 4 for one of the samples stressed at 100°C.

The increased rate of relaxation may be attributed to the loss of practically all of the antioxidant found in the original SBR. The extracted sample had also hardened

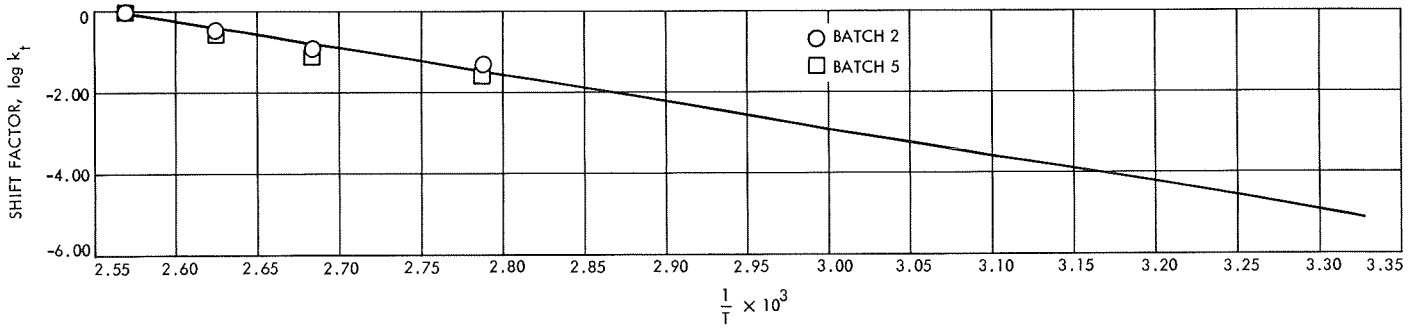


Fig. 3. Temperature dependence of shift factor, log K_t

Table 6. Test time at which stress relaxation is due to viscoelastic behavior

Temperature, °C	Time, s
85	192
100	38
108	15
115	6

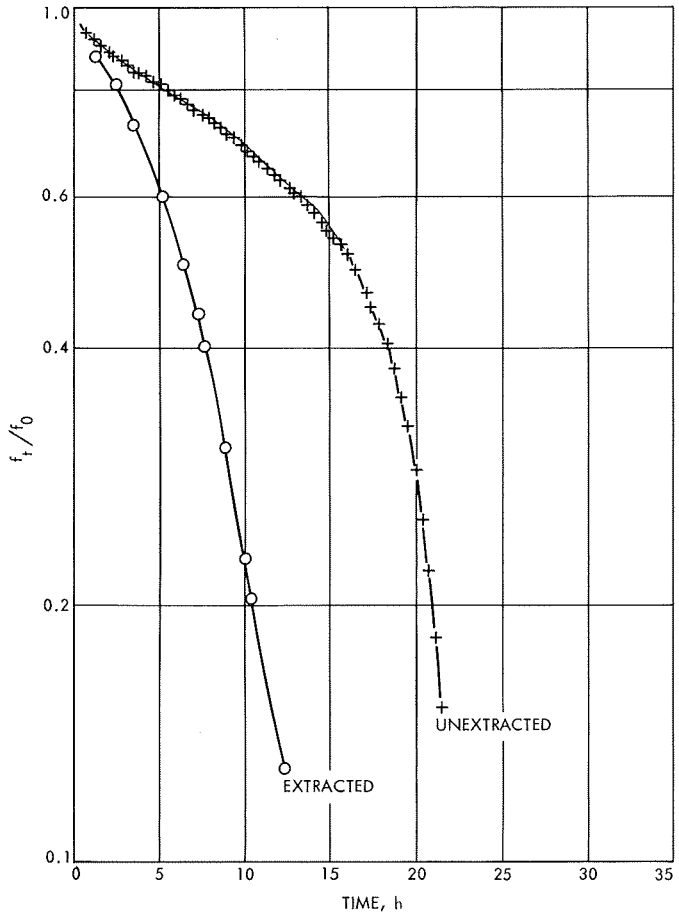


Fig. 4. Stress relaxation curves at 100°C for extracted and unextracted SBR (batch 2)

considerably, indicating extensive crosslinking by oxidation. Hardening and crazing were also more predominant in unextracted samples stressed at 108 and 115°C.

b. Experiments in dry nitrogen atmosphere. Stress relaxation experiments carried out in nitrogen showed that, regardless of the temperatures used, which ranged from 43 to 115°C, the decay in the tensile force was less than 5% after 85–90 h. Also, in this atmosphere there was no observable change in the physical state of the stressed specimens, such as hardening, crazing, or opacity observed in samples stressed in air. This indicates that relaxation due to chemical changes is caused by atmospheric oxygen, rather than by thermal effects, at the time and temperature range used.

References

1. Landel, R. F., "Office of Naval Research Structural Mechanics Series, Mechanics and Chemistry of Solid Propellants," in *Proceedings of the Fourth Symposium on Naval Structural Mechanics*, p. 575, Pergamon Press, New York, 1967.
2. Fedors, R. F., and Landel, R. F., "A Test of the Predictability of the Properties of Filled Systems," presented at the AIAA Meeting, Washington, D.C., June 1967.
3. Tobolsky, A. V., Prettyman, I. B., and Dillon, J. H., "Stress Relaxation of Natural and Synthetic Rubber Stocks," *J. Appl. Phys.*, Vol. 15, p. 380, 1944.
4. Thirion, P., and Chasset, R., "Relative Contributions of Viscoelasticity and Aging to the Relaxation of Rubber Vulcanizates," *Rubber Chem. Technol.*, Vol. 36, p. 50, 1963.
5. *The Chemistry and Physics of Rubber-Like Substances*, Chap. 18. Edited by L. Bateman. John Wiley & Sons, Inc., New York, 1963.
6. Tobolsky, A. V., *Properties and Structure of Polymers*, p. 226. John Wiley & Sons, Inc., New York, 1960.

C. Viscoelastic Behavior of Elastomers Undergoing Scission Reactions, *J. Moacanin, J. J. Aklonis,[2] and R. F. Landel*

Previously (SPS 37-57, Vol. III, pp. 169–172, and SPS 37-58, Vol. III, pp. 199–201), a mathematical formulation was evolved that describes the behavior of elastomers simultaneously experiencing mechanical relaxation and scission reactions. In this article, an analysis is presented of the parameters which characterize the viscoelastic relaxation of an elastomer. In presence of scission reactions, these parameters are not constant but become time-dependent in a manner determined by the kinetics of the chemical reactions.

Plazek (Ref. 1) showed that, in addition to the usual time–temperature reduction, crosslink density ν is also a reduction variable for elastomers such as natural rubber or SBR. Here we examine the feasibility of representing Plazek's empirical reduced master curve with the two-parameter generating function used previously (SPS 37-58, Vol. III, pp. 199–201). For if one can analytically represent the master curve of an elastomer for a given crosslink density and temperature, then a convenient

[2]Department of Chemistry, University of Southern California, Los Angeles, California.

mathematical formulation is provided to describe the behavior of the elastomer at any other temperature or crosslink density.

In the analysis of his experimental results, Plazek employed a terminal retardation function $\psi_\nu(t)$ defined in terms of reduced shear creep compliance $J_\nu(t/a_\nu)$ as

$$\log J_\nu\left(\frac{t}{a_\nu}\right) = \log J_e(\nu) + \log\left[\frac{J_e(\nu_{ref})}{J_e(\nu)}\right] + \log \psi_\nu\left(\frac{t}{a_\nu}\right) \tag{1}$$

where the second term on the right-hand side shifts vertically the equilibrium compliance J_e to the level of the reference state. This shift is linear in crosslink density and presents no appreciable problems. The scale factor a_ν provides the reduced time scale and is an exponential function of ν. It should be noticed that

$$\lim_{t\to\infty} \psi\left(\frac{t}{a_\nu}\right) = 1 \tag{2}$$

while at finite times $\psi(t/a_\nu) < 1$. Values of $\psi(t/a_\nu)$, a function independent of crosslink density, have been measured by Plazek and are listed in Table 7. (Note: Here expressions are written in terms of ν, whereas Plazek uses molecular weight between crosslinks M_c with $\nu \sim 1/M_c$.)

Table 7. Normalized terminal retardation function $\psi(t/a_\nu)$ according to Eq. (3)

$\log (t/a_\nu)$	$-\log \psi\ (t/a_\nu)$		
	Experimental (Ref. 1)	Calculated	
		$\tau = 0.01$, $m = 0.118$	$\tau = 100$, $m = 0.17$
−1	0.2315	0.246	0.626
0	0.1906	0.199	0.503
1	0.1578	0.159	0.394
2	0.1286	0.126	0.301
3	0.1039	0.099	0.224
4	0.0821	0.078	0.163
5	0.0606	0.0604	0.117
6	0.0472	0.0468	0.082
7	0.0355	0.0361	0.0573
8	0.0277	0.0278	0.0396
9	0.0214	0.0213	0.0271
10	0.0163	0.0163	0.0185
11	0.0122	0.0125	0.0127
12	0.0086	0.0095	0.0086
13	0.0058	0.0073	0.0058
14	0.0040	0.0055	0.0039

The two-parameter function consistent with our previous function for the relaxation modulus $G(t)$ (Eq. 5, SPS 37-58, Vol. III, pp. 199–201) gives

$$\psi\left(\frac{t}{a_\nu}\right) = 1 + \left(\frac{t}{a_\nu \tau}\right)^{-m} \tag{3}$$

where m and τ are chosen to give the best fit for the experimental data. Values of $m = 0.113$ and $\tau = 0.01$ were found to give an excellent fit for the experimentally determined creep curve. Calculated values of $\psi(t/a_\nu)$ are listed in Table 7; these values are within 5% of the experimentally determined values over at least 12 decades of log-time. It should be noticed, however, that deviations from observed behavior increase progressively at both the long- and short-time ends of the curve. But since we are concerned with behavior at long times, it is advantageous to optimize the long-time portion of the curve at the expense of that for short times. The last column of Table 7 indicates that such a fit is possible with $m = 0.17$ and $\tau = 100$.

Close scrutiny of Plazek's results reveals that there may be a slight discrepancy between the long-time end of the reduced master curve and the experimental data for the sample which was used to determine the end portion of the curve (sample J from Ref. 1). The curve for sample J does not flatten out as fast as the master curve indicates. Table 8 lists values of $\psi(t)$ for this sample measured at 30°C. [$J(\infty)$ was assumed to be 2.0×10^{-7}

Table 8. Normalized terminal retardation function $\psi(t)$ for sample J (Ref. 1)

$\log t$	$-\log \psi\ (t)$	
	Experimental (Ref. 1)	Calculated
		$\tau = 10^{-18}$, $m = 0.0632$
−1	0.0350	0.0351
0	0.0315	0.0305
1	0.0252	0.0265
2	0.0223	0.0230
3	0.0200	0.0200
4	0.0173	0.0173
5	0.0150	0.0150

cm^2/dyne.] The choice of parameters to fit this curve and resultant values of $\psi(t)$ are listed in Table 8.

This analysis has led us to the conclusion that Eq. (3) will serve as an excellent generating function for $\psi(t/a_v)$; substitution of this function into Eq. (2), along with some additional information pertaining to the elastomer crosslink density, will permit an accurate calculation of the relaxation behavior of non-reacting networks at any temperature and any crosslink density.

As mentioned previously, we have been able to find only one set of experimental results (Ref. 2) which appeared to be suitable to test our mathematical formulation. Detailed examination of these data, as reported in Ref. 2, has indicated that they do not satisfy the principle of time–temperature superposition. Compliance with this empirical principle is a prerequisite to the application of the techniques discussed in this and previous articles. Since adherence to the principle of time–temperature reduction has been rather well substantiated by many investigators, it appears that there are inconsistencies in the data and thus the data are unsuitable for our analysis. Thirion and Chasset (Ref. 2) were interested only in a semi-quantitative way in this portion of their data, whereas we analyzed the same data with a very sensitive method. Thus, the accuracy of these data is insufficient for this work, although the discrepancies amount to only a few percent.

References

1. Plazek, D. J., *J. Poly. Sci.*, Vol. A-2, No. 4, p. 745, 1966.
2. Thirion, P., and Chasset, R., *Rubber Chem. and Tech.*, Vol. 37, p. 617, 1964.

D. Investigations on Sterilizable Battery Separators, *E. F. Cuddihy, D. E. Walmsley, and J. Moacanin*

1. Introduction

Sterilizable battery separator membranes have been prepared by grafting poly(acrylic acid) onto thin films of polyethylene and then converting the acrylic acid to the potassium salt. This article is another report on the continuing studies (SPS 37-50, -52, -54, -56, Vol. III) on the fundamental chemical and physical properties of this membrane system.

2. Silver Oxide–Separator Interaction

The graft copolymer films are used as separators in silver–zinc secondary batteries to prevent soluble silver provided by the silver electrode (50 mg/liter Ag^+ ion in battery electrolyte) from reaching the zinc electrode. Silver ion is reduced to form on the zinc electrode deposits of metallic silver which would cause local cell action and failure of the cell.

During use in the battery, the initially clear membranes become progressively blacker due to the deposition within the membrane of precipitated silver products. As previously shown (SPS 37-56, Vol. III, pp. 186–188), the electrical resistivity of the membrane continuously decreases with increasing deposition, a condition which could lead to the development of an internal short circuit as the membrane is pressed between the two electrodes.

A detailed analysis of the interaction between the membranes and soluble silver oxide in battery electrolyte is now underway. Initial emphasis has stressed the development of analytical methods for the qualitative and quantitative determination of the silver products. Analytical methods for acrylate content, polyethylene crystallinity, and other important parameters are detailed in the previous SPS contributions.

a. Qualitative analysis. The elemental constituents of a blackened battery separator were identified by using an electron microprobe. This device scans the sample with a stream of electrons sufficiently energetic to excite each of the elements in the sample and cause them to emit their individually characteristic emission spectra. Detectors set to match a particular element's spectra act as a go, no-go indicator, indicating either the presence or absence of a given element. Since the intensity of the emission is proportional to concentration, relative differences in concentration from sample to sample can be measured but absolute determinations are generally not accurate.

Electron microprobe analysis of a membrane blackened by exposure to a silver-oxide-saturated 40% KOH–H_2O solution was found to contain only silver, oxygen, carbon, hydrogen, and potassium. Also, an analysis after battery service revealed identical elements with the only exception that calcium had replaced potassium as the acrylate counterion (for a detailed discussion see SPS 37-54, Vol. III, pp. 100–104). Soaking the membranes in 5% HCl effected removal of the acrylate counterion as well as causing the color of the membranes to change from black to light tan. It was suspected that the deposited silver oxide, which is black, reacted with HCl to form white silver chloride, which would lighten the color of the membrane. Microprobe analysis verified the removal

of the acrylate counterion, the reduction in oxygen, and the presence of chlorine.

But microprobe analysis cannot distinguish between an element chemically combined or valently neutral; hence, this method cannot provide a quantitative breakdown on the relative amounts of metallic silver and silver oxide. Nevertheless, it provided the first experimental verification that the silver products were some combination of metallic silver and oxides of silver, there being no other silver salts such as nitrates, phosphates, sulfates, etc.

b. Quantitative analysis (Ref. 1). Potassium cyanide (KCN) solutions will dissolve silver salts, but not metallic silver unless oxygen (or air) is bubbled through the solution. Exposing a blackened membrane to oxygen-free KCN, solutions will result in removal of the silver salts. After constant weight is achieved, the metallic silver can then be removed by bubbling oxygen (or air) through the KCN solution. Alternatively, one-half of a sample can be exposed to an oxygen-free KCN solution to remove only silver salts with the other half exposed to an oxygen-containing KCN solution to remove all silver products. Weight measurements taken before and after permit calculation of the metallic silver and silver oxide concentrations.

c. Preliminary deposition study. A sample of battery separator was exposed for 24 days at 85°C to a silver-oxide-saturated 40% KOH–H_2O solution. The concentration of poly(potassium acrylate) dropped in that time period from an initial 27 wt % to a final 7 wt % with a corresponding pick up of 19.3 wt % silver products, of which 15.2 wt % was metallic silver and 4.1 wt % was silver oxide. These results correspond to a loss of about one acrylate group from the membrane for each atom of deposited neutral silver confirming the occurrence of oxidative degradation reactions. Also, it appears that the base polyethylene film retains its structure so that the battery separator becomes less hydrophilic, resulting in degradation of battery performance.

Continued studies will monitor the rate of silver pickup and acrylate loss as a function of time and temperature to permit estimates of lifetime under service conditions.

3. DOW 400

Dow 400 is a thin-film polyethylene (Dow Chemical Company) which is currently being investigated for use in the battery separator membrane. Prior to this, Petrothane 301 (Phillips Petroleum Company) had been used exclusively. A comparison of the properties of those two thin polyethylene films is given in Table 9.

Both films have identical crystallinities (and hence densities) which mainly control the film toughness. The apparent improvement in toughness of Dow 400 over Petrothane 301 is due to the former's larger film thickness rather than any changes in the nature of the polyethylene.

The key differences between the two is unsaturation, Dow 400 being totally saturated whereas Petrothane 301 contains 0.12 mM/g of pendant vinylidene groups. Although double bonds in the polyethylene could influence the radiation grafting reaction, the acrylic acid content of a membrane prepared using Dow 400 (GX-4-362) was comparable to that for the Petrothane 301 systems. The density and length of grafted chains may have been different, but unfortunately no method has yet been devised for quantitative measurements.

Table 9. Comparison of polyethylene films[a]

Parameter	Petrothane 301	DOW 400
Density at 22°C, g/cm³	0.925	0.926
Melting point, °C	107	107
Crystallinity, %	50	50
Degree of branching, pendant groups per 100°C atoms	1.7 to 2.0	2.3
Unsaturation, mM/g of pendant vinylidene groups	0.12	No unsaturation
Pendant groups	CH_3 and vinylidene groups	Only CH_3 found
Weight average molecular weight, M_w	48,200	73,600
Number average molecular weight, M_n	15,000	31,300
M_w/M_n	3.18	2.35
Additives[b]	Detected but unidentified	None detected
Solubility, % in α-chloro-naphthalene (>107°C)	100	100
Thickness, μm	25	40

[a]See SPS 37-50, Vol. III, p. 167 for description of test methods.
[b]See SPS 37-52, Vol. III, p. 98 for details concerning Petrothane 301.

4. Hydration of Calcium Salts

It has been previously reported that potassium counterions in the graft copolymer membrane caused strong sorption of atmospheric moisture. This water is not strongly bound to the potassium ion but is dissolved throughout the poly(potassium acrylate) phase, thus plasticizing the film. On the other hand, membranes containing calcium do not sorb water and, therefore, are brittle. However, from variations in weight of calcium-containing membranes being dried under high vacuum, it has been discovered that calcium absorbs two waters of hydration as strongly bound water. This water, strongly held to the calcium ion, is not free to plasticize the calcium acrylate. It exists in a stoechiometric amount at all levels of humidity, except under high vacuum where it requires several hours at room temperature to be removed.

The elimination of these two waters of hydration results in the disappearance of a strong IR peak at 3 μm (Fig. 5), a peak which is also observed for the potassium-containing membranes.

5. Variations in Acrylic Acid Concentration

The acrylic-acid concentration of the membranes is now monitored on all samples under investigation, using a previously developed method (SPS 37-54, Vol. III, pp. 100–104). Usually, the concentrations vary between 20 and 30 wt %, both for different batches of membranes as well as within a given batch. However, certain areas of the membranes have been found to have concentration ranging as high as 35 wt % to a surprisingly low 4 wt %. These extremes may make difficult meaningful comparison of performance among batteries employing different swatches of battery separator.

6. GX-115

Grafting of acrylic acid onto polyethylene is initiated by exposure to a cobalt 60 source and the heat therefrom

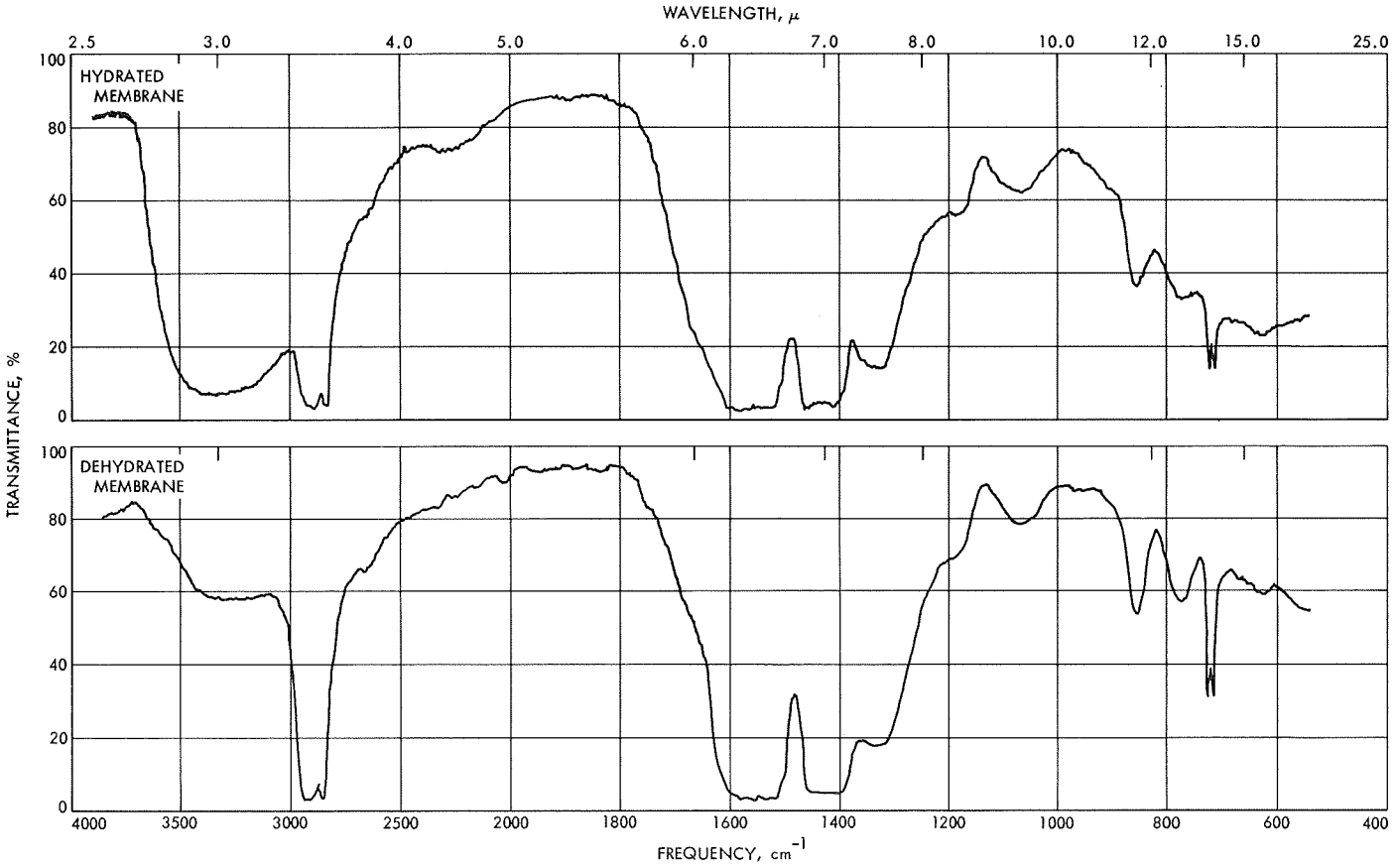


Fig. 5. IR traces of battery separator membranes containing calcium counterions

introduced into the system is continuously carried away by a cooling stage to prevent an undesirable temperature climb. The higher temperatures are to be avoided because this favors formation of large quantities of a poly(acrylic acid) homopolymer which plugs up the reaction vessel and makes very tedious the removal of the grafted membrane. This occurred during the grafting reaction of batch GX-115 where presumably the cooling unit failed to function.

In assessing battery performance, it was observed that cells prepared using GX-115 separators had in general noticeably improved performance over cells prepared with other membranes. It was also observed that the GX-115 separator was considerably more opaque than the usually clear membranes. These observations prompted a detailed investigation of this system.

So far, it has been determined that the crystallinity content of the polyethylene (PE) phase was unchanged after grafting. Hence, the increase in temperature was not enough to effect measurable melting and again for this system all of the grafting has occurred in the amorphous phase initially present in the polyethylene film. The acrylic acid concentration appears to be 25 to 26 wt %, but one section of the membrane had 35 wt %.

Differential scanning calorimetry (DSC) was employed (SPS 37-50, Vol. III, pp. 166–169) to measure the PE crystallinity of the GX-115 films. Previously, it was observed that in the presence of potassium counterions, two DSC endotherms are observed, one at 107°C for the melting of the crystalline PE and another starting at 137°C for the dehydration of the loosely sorbed water. Typical curves are shown in Fig. 6. The area under the PE melting peak can be used to calculate the crystallinity content. If calcium replaces potassium as the counterion, there is no endotherm at 137°C due to the absence of loosely sorbed water, but this does not affect the PE melting peak.

DSC curves for GX-115 membranes containing potassium and calcium counterions are shown in Fig. 7. For the potassium-containing membrane, an unexpected shoulder occurs on the low-temperature side of the PE melting endotherm. In addition, the area under the curve predicts that the crystallinity of the PE in the grafted film is in excess of 80% as compared to the initial 50%, an obviously erroneous result. It can only be supposed that the shoulder is the result of the superposition of a second endothermic process occurring in the same temperature range as the crystalline melting of polyethylene. The DSC curve for the calcium-containing membrane is normal, having no shoulder and an area under the curve which predicts a proper PE crystallinity of near 50%. Additionally, the higher temperature dehydration endotherms for the potassium-containing membrane were greatly reduced in intensity (compare with Fig. 6).

These results suggested that the dehydration from the GX-115 potassium-containing membranes was occurring at a much lower temperature than that for other batches. However, a run made on a DSC sample completely dehydrated under high vacuum at room temperature resulted in a shoulder and no higher temperature endotherm. Also, as is general practice, the sample was weighed before and after the DSC run and there was no weight change, indicating no loss of water from

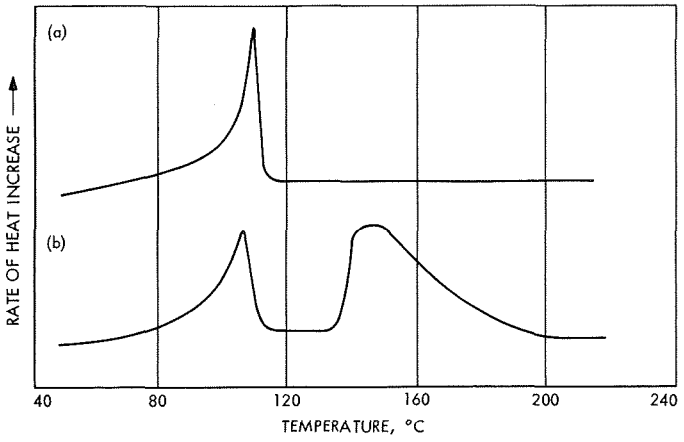


Fig. 6. Differential scanning calorimetric endotherms for (a) 1.8-mg polyethylene film, Petrothane 301, 50% crystallinity, and for (b) 3-mg graft copolymer of polyethylene and poly(potassium acrylate)

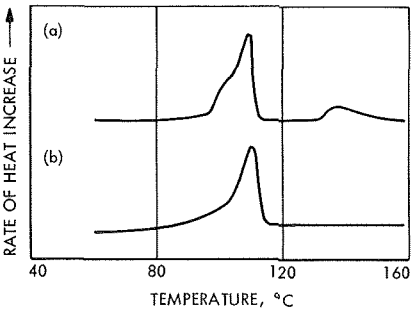


Fig. 7. Differential scanning calorimetric endotherms for GX-115 membranes (a) 3.9-mg sample with potassium counterion, (b) 8.2-mg sample with calcium counterion

dehydration. Still the shoulder appeared. This effect and its possibility of being a significant clue to understanding the GX-115 battery separator is still under investigation.

Reference

1. Latimer, W., and Hildebrand, J., *Reference Book of Inorganic Chemistry*, 3rd Edition. The Macmillan Company, New York, 1951.

XIX. Research and Advanced Concepts

PROPULSION DIVISION

A. Performance and Stability of a Hollow Cathode Ion Thruster, *E. V. Pawlik*

1. Introduction

A program is presently under way to study the system aspects of solar-electric propulsion. Some system results have been presented in Refs. 1 and 2 for an ion thruster that employs an oxide cathode. The lifetime and performance limitations of the thruster using this type of cathode have prompted its replacement with a hollow cathode version.

2. Experimental Setup

An ion thruster was operated in a 3- × 7-ft vacuum tank at pressures ranging from 3×10^{-6} to 9×10^{-6} torr. The ion thruster utilized mercury as the propellant and provided a nominal 20-cm-diam ion beam. It is, with some small modifications, the same thruster as described in Ref. 1 and SPS 37-51, Vol. III, pp. 124–128. These modifications included cathode replacement with a hollow cathode, pole piece alterations, and a decrease in acceleration distance. Two vaporizers were used. The spacing between ion accelerating grids was reduced from 0.175 cm to 0.125 cm. A hollow cathode with a 0.063-cm-diam orifice was used and was fabricated with a 0.465-cm-diam fin on the emitting surface to provide radiation cooling. This design has been operated for about 100 h of accumulated short term testing with no wear observed at the orifice. Details of the cathode and pole piece location are shown in Fig. 1. Instrumentation to monitor thruster performance included panel meters, slow-speed Leeds and Northrup strip chart recorders with a response time on the order of 160 ms and high-speed Brush strip charts with a response time on the order of 2 ms. The slow-speed recorders indicated when constant thruster performance was obtained with respect to long term thermal time constants of the thruster, vaporizers, and connecting propellant lines. Meaningful thruster plasma stability measurements could only be obtained with the high-speed recorder.

3. Test Results

Several thruster pole piece configurations had been tested prior to the present design. In general, stable operation of the plasma discharge could be obtained only above some minimum arc current which at times was

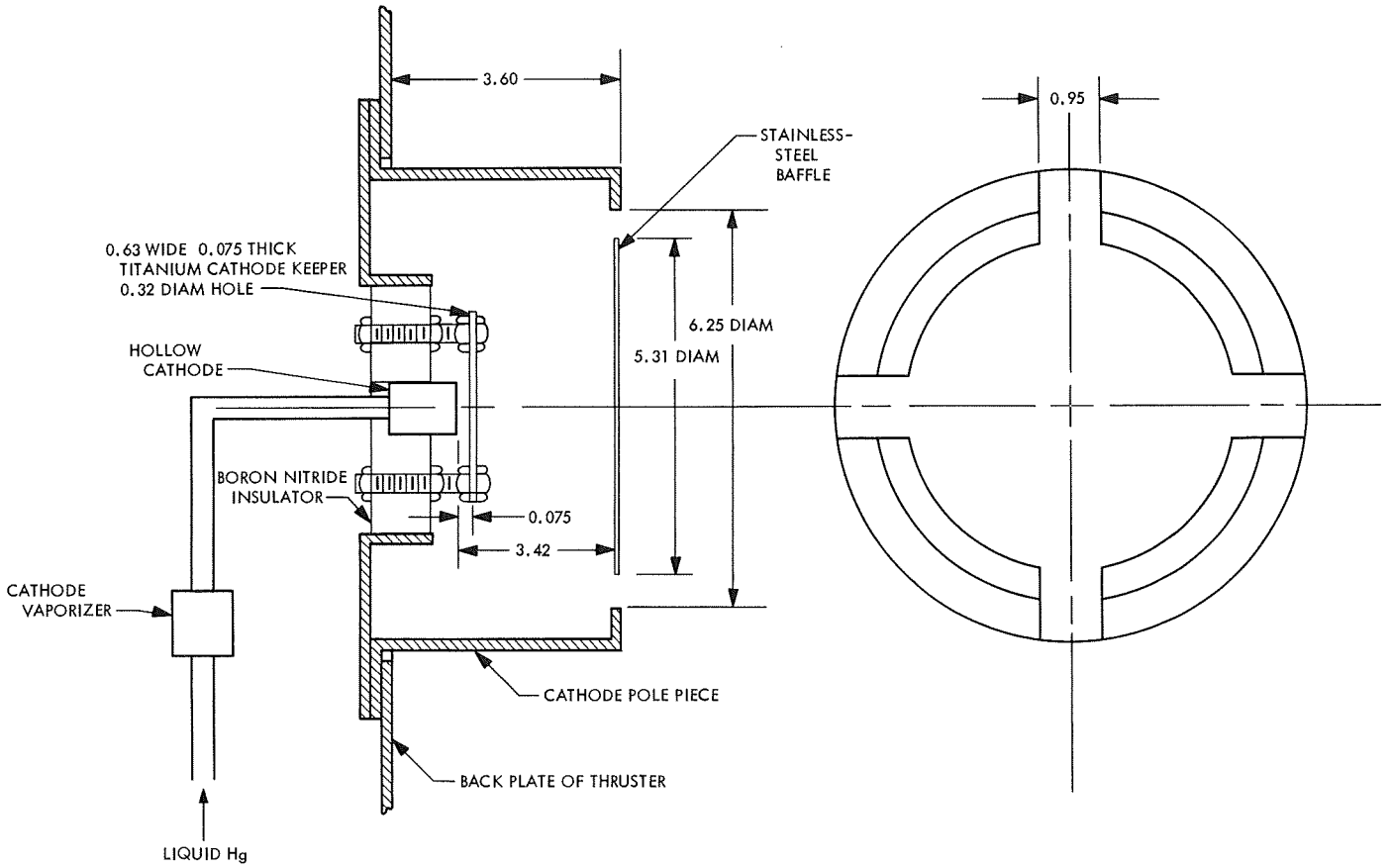


Fig. 1. Cathode construction and mounting

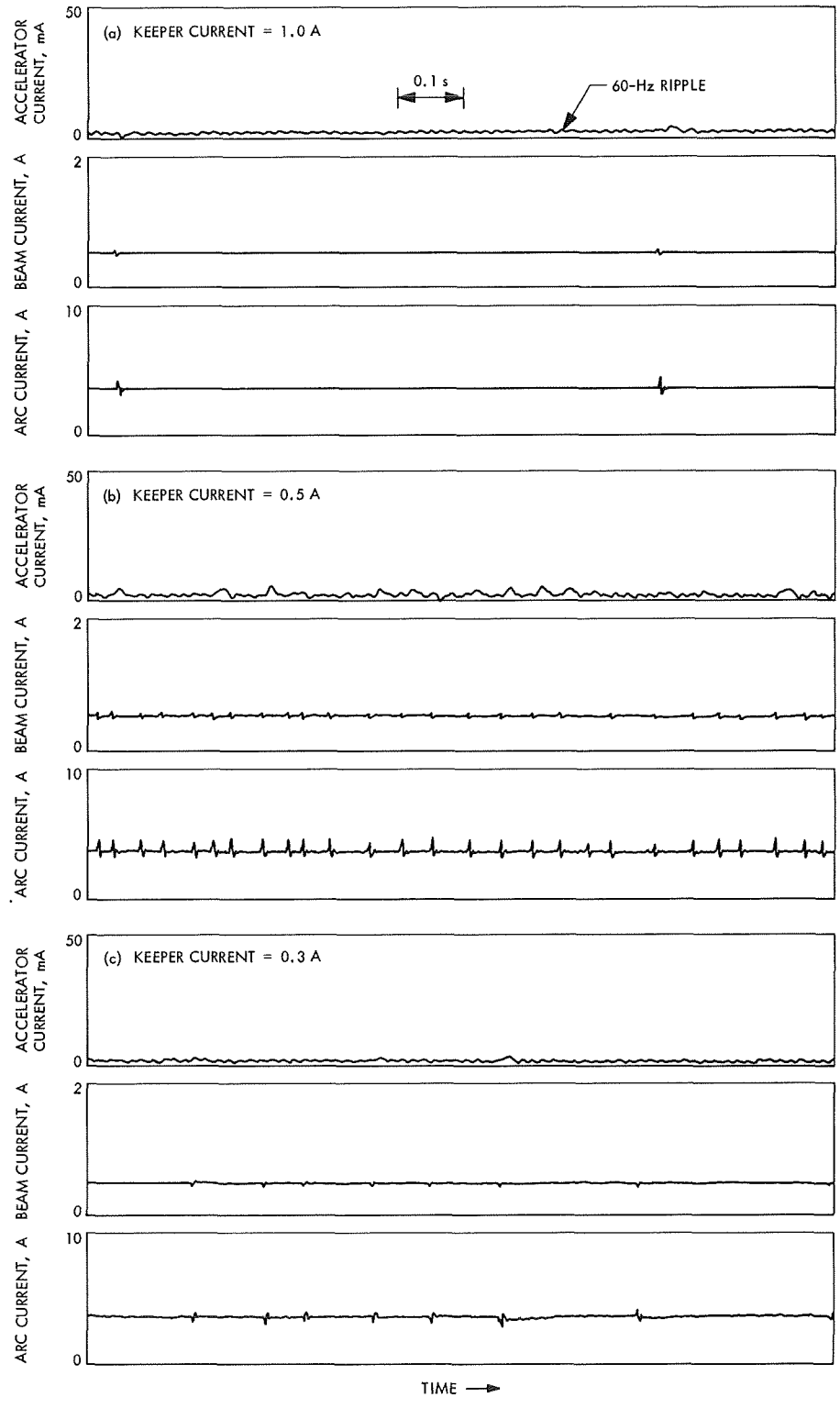


Fig. 2. Thruster current traces for different values of cathode keeper current

close to high propellant utilization (percentage input propellant ionized and accelerated as the ion beam) operation. In some regions of operation the arc and accelerator currents were observed to contain high noise levels. The dc levels of operation were obscured by these noise levels in some cases. In the present design the noise levels and minimum arc currents to permit a plasma discharge were considerably lower and therefore a reduction of this noise content was attempted.

Traces of the arc, beam, and accelerator currents are presented in Fig. 2 for three different values of cathode keeper current. Arc current "noise" was present as sudden, nearly periodic changes in the arc current level. The restoration and ringing following these surges are believed to be the reaction of the arc power supply which was operated in a current-limited mode. The frequency of occurrence and the amplitude of these noise spikes are presented in Fig. 3 as a function of the keeper current. This noise was minimized at keeper currents of either zero or at 1.2 A and above.

The effect of magnet current on the arc current noise is presented in Fig. 4. Noise-free operation was obtained at a magnet current of 0.405 A. Decreasing noise amplitude at higher magnetic fields suggests a higher stable region. The instrumentation is attenuating in this region however, and therefore operation in this area should be checked with oscilloscope measurements.

The arc current noise content was minimized as the thruster was operated at a low (4.1 g/h) propellant flow rate and near propellant utilization of 90%. The 0.405-A magnet and 1.2-A keeper currents minimized the arc chamber noise content in this region, and these values were maintained constant during thruster mapping. Performance results are presented in Fig. 5 for thruster operation over the entire 2:1 range of output power normally of interest for power matching to a solar array with a varying solar constant. The cathode flow was adjusted to maintain the arc voltage near 35 V at 90% propellant utilization for each flow rate. During this mapping the low response instrumentation indicated noise-free operation with the exception of some high propellant utilization points where some accelerator current instabilities were evident. Each data point was checked on the high-speed strip chart. Most indicated quiet operation. Two points contained some minor noise

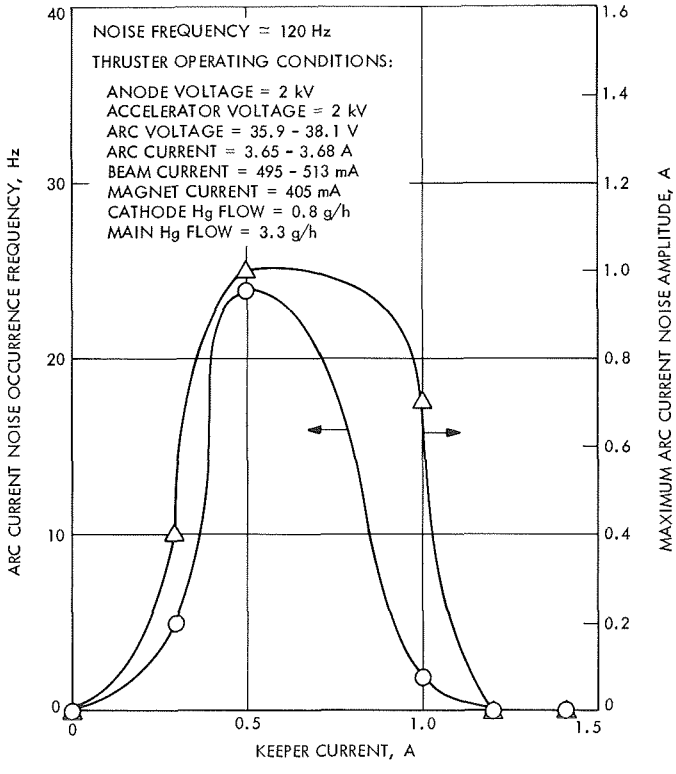


Fig. 3. Arc current noise signals as a function of keeper current

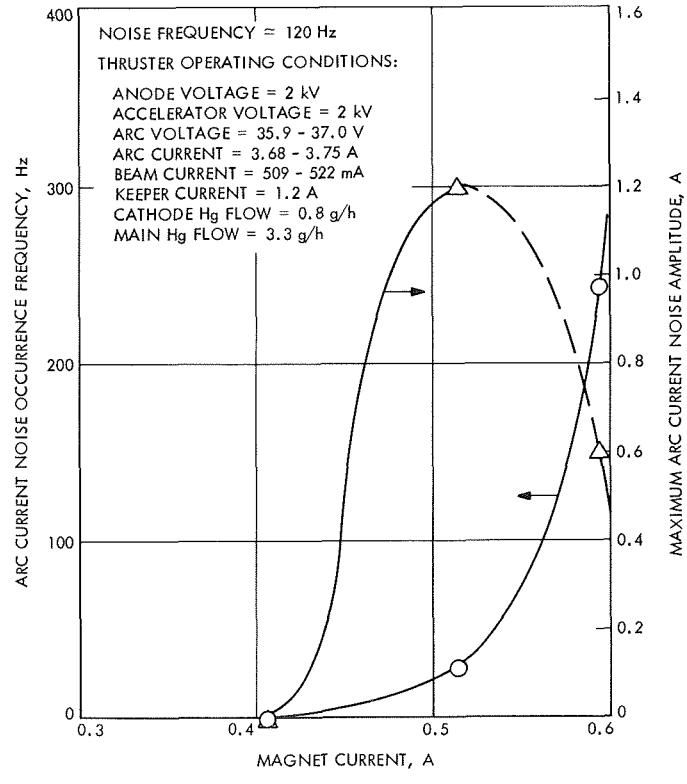


Fig. 4. Arc current noise signals as a function of magnet current

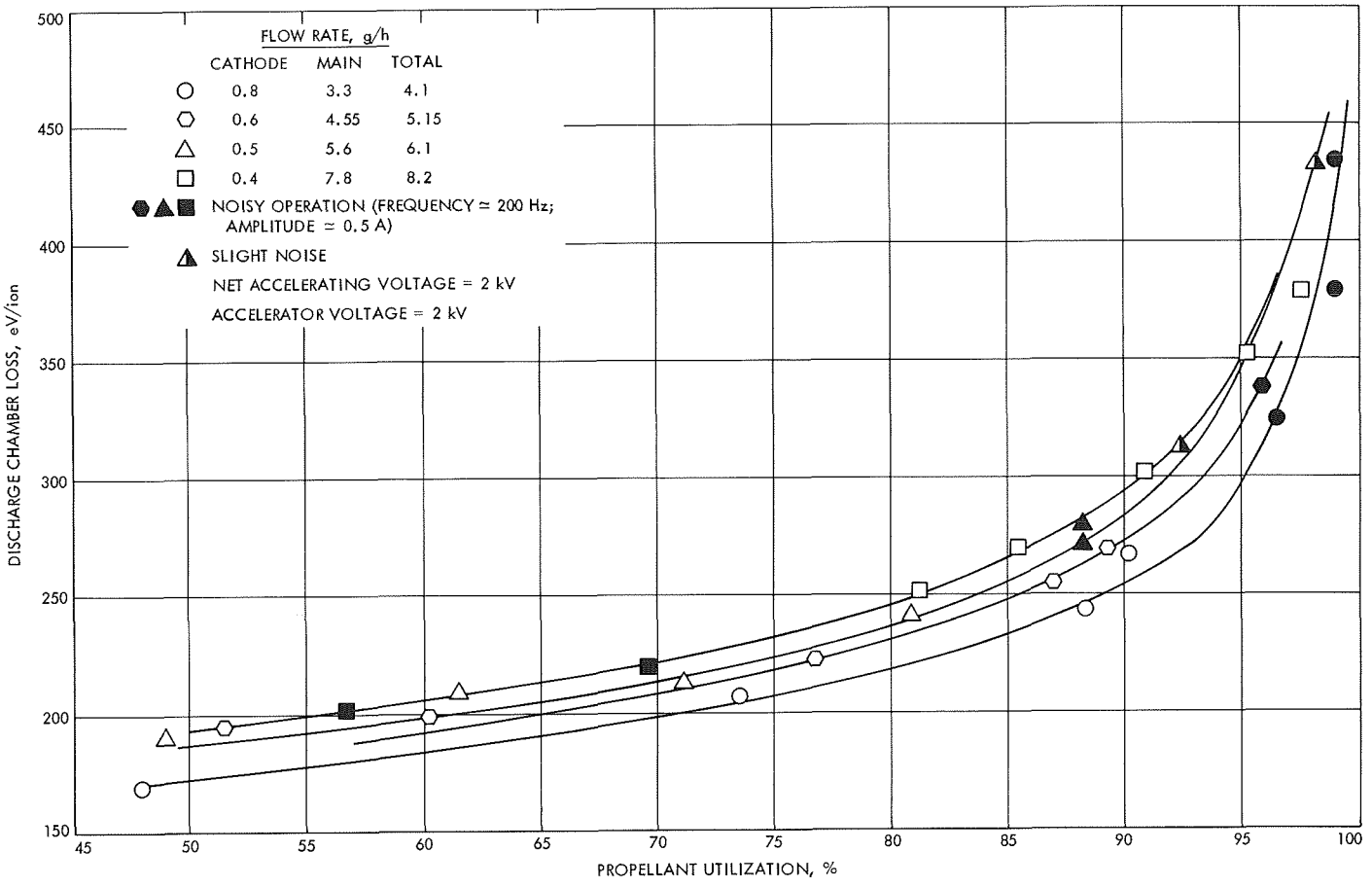


Fig. 5. Effect of propellant flow rate on discharge loss

levels and are indicated by half-solid data points. The solid data points indicate where noise was observed on the arc current. At these points continuous oscillations of the arc current at a frequency of 200 Hz and an amplitude near ±0.5 A were observed. This frequency is close to the ion cyclotron frequency (about 170 Hz) for the nominal 20-gauss field that was employed. Tests were not run to examine the effects of magnetic field configuration, mercury propellant flow distribution, pole piece geometry, or cathode temperature on the arc current noise content. It would appear to be desirable to eliminate the noise at least in the neighborhood of 90% propellant utilization, which is the normal thruster operating region. The effects of these perturbations on power conditioner control circuitry should also be examined.

The accelerator current obtained during the above mapping is presented in Fig. 6. The close grid spacing produced a lower percentage of accelerator to beam current than had been obtained during prior testing (Ref. 1) with these grids at the larger separations. This suggested that the interception at 100% utilization is at least partially comprised of direct ion beam impingement.

Performance of the thruster for the data points obtained that were closest to 90% propellant utilization are presented in Table 1. Throttling at this constant utilization appears to present no significant penalties in power matching such as were reported in Ref. 3.

4. Conclusions

(1) Use of high-speed instrumentation greatly expedited the adjustment of thruster parameters for noise-free operation over a wide range of propellant utilization and output power.

(2) Anomalies in arc chamber operation exist and remain to be resolved.

(3) Throttling of a hollow cathode thruster at a constant propellant utilization over a 2:1 range of

Table 1. Hollow cathode ion thruster performance data

Mercury propellant flow rate, equiv amperes	Arc voltage, V	Beam power, W	Power loss						Power efficiency	Propellant utilization efficiency		Total efficiency	
			Arc chamber, W	Accelerator, W	Cathode, W	Magnet-manifold, W	Main vaporizer, W	Neutralizer,[a] W		Excluding neutralizer	Including neutralizer estimate[b]	Excluding neutralizer flow	Including neutralizer estimate[b]
1092	36.5	1986	300	15.0	24	3.8	12.8	20	84.1	90.9	90.0	76.4	75.7
814	36.8	1440	196	12.0	24	3.8	11.9	20	84.3	88.4	87.2	74.5	73.5
690	38.1	1230	166	10.0	24	3.8	11.6	20	83.9	89.2	87.6	74.8	74.3
547	34.5	988	126	7.6	24	3.8	11.4	20	83.7	90.2	88.4	75.5	74.0

[a]Cesium neutralizer operation (SPS 37-58, Vol. III, pp. 212–215).

[b]Including cesium neutralizer flow.

output power can be achieved with small penalties in total thruster operating efficiency.

References

1. Pawlik, E. V., Macie, T. W., and Ferrera, J. D., "Electric Propulsion Systems Performance Evaluation," Paper 69-236, presented at the Seventh AIAA Electric Propulsion Conference, Williamsburg, Va., Mar. 1969.
2. Macie, T. W., Pawlik, E. V., Ferrera, J. D., and Costogue, E. N., "Solar-Electric Propulsion System Evaluation," Paper 69-498, presented at the Fifth AIAA Propulsion Joint Specialist Conference, U.S. Air Force Academy, Colorado Springs, Colo., June 9–13, 1969.
3. Bechtel, R. T., "Performance and Control of a 30-Centimeter Diameter Low Impulse Kaufman Thruster," Paper 69-238, presented at the Seventh AIAA Electric Propulsion Conference, Williamsburg, Va., Mar. 1969.

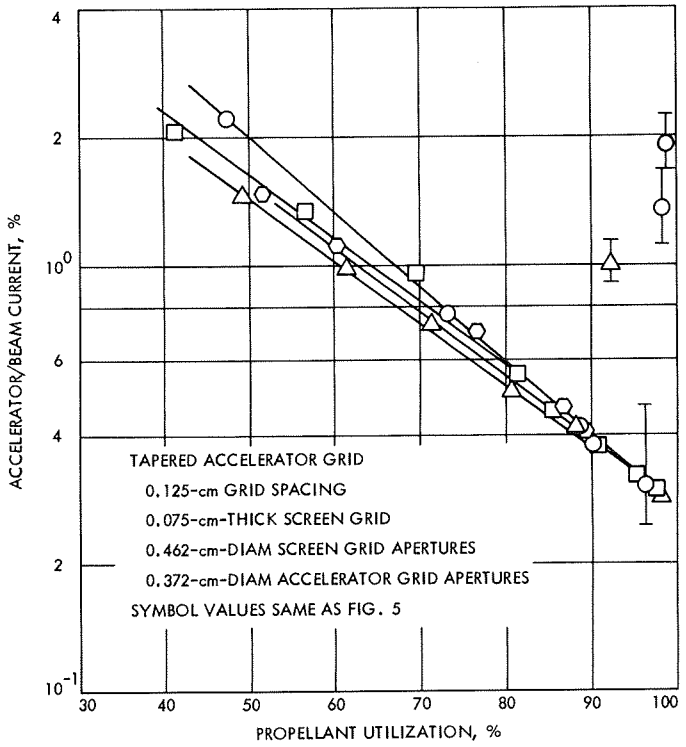
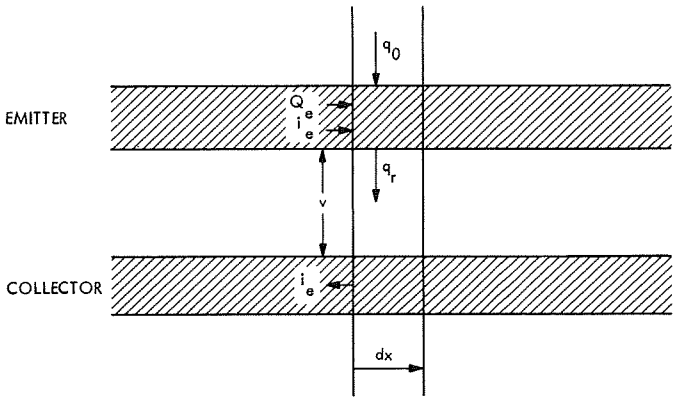


Fig. 6. Effect of propellant flow rate on accelerator impingement

B. One-Dimensional Diode Heat Transfer Program Coupled to Thermionic Performance, *J. P. Davis, M. Diethelm, and N. K. Simon*

1. Introduction

Electrical and heat balances are taken for a differential segment of diode length for the case where emitter and collector current flows are countercurrent (see Table 2 for definition of symbols used):



$$\frac{dv}{dx} = -\, i_e(x)\, \frac{\rho_{\text{eff}}}{A_e} \qquad (1)$$

Table 2. Nomenclature

A_c	= collector cross-sectional area, cm^2
A_e	= emitter cross-sectional area, cm^2
C_e	= emitter circumference, cm
F	= specified fraction of total heat input lost down emitter lead, $0 \leq F \leq 1$
$i(x_f)$	= current at end of diode, A
j	= local current density, A/cm^2
k	= emitter conductivity, W/cm-°K
k_1	= lead conductivity, W/cm-°K
p	= average diode power density measured at lead end, W/cm^2
Q	= $-\, kA_e\, (dT/dx)_{x=x_f}$, W
Q_e	= local axial emitter heat flow, W
q_0	= local heat input to emitter, W/cm^2
q_r	= local heat transferred across inter-electrode gap, W/cm^2
T	= temperature along diode axis, °K
T_s	= temperature at end of lead, °K
v	= local voltage difference between emitter and collector, V
v_{net}	= voltage at end of lead, V
x	= distance along diode axis, cm
y	= ratio of length of lead to cross-sectional area of lead, cm^{-1}
ρ_c	= collector resistance, Ω-cm
ρ_e	= emitter resistivity, Ω-cm
ρ_{eff}	= effective resistivity, Ω-cm
ρ_1	= electrical resistivity of lead, Ω-cm

where

$$\rho_{\text{eff}} = \rho_e + \rho_c \frac{A_e}{A_c}$$

$$i_e(x + dx) = i_e(x) + j(v, T)\, C_e\, dx \tag{2}$$

from which

$$\frac{di_e}{dx} = j(v, T)\, C_e \tag{3}$$

Differentiating Eq. (1), substituting in Eq. (3), and rearranging gives the voltage–current density relation

$$\frac{d^2v}{dx^2} = -\frac{\rho_{\text{eff}} C_e}{A_e} j(v, T) \tag{4}$$

The emitter heat balance:

$$Q_e(x) = -kA_e \frac{dT}{dx} \tag{5}$$

$$Q_e(x + dx) = Q_e(x) + q_0(x)\, C_e\, dx - q_r(v, T)\, C_e\, dx + i_e(x)^2 \frac{\rho_e}{A_e} dx \tag{6}$$

from which

$$\frac{dQ_e}{dx} = q_0(x)\, C_e - q_r(v, T)\, C_e + i_e^2(x) \frac{\rho_e}{A_e} \tag{7}$$

Differentiating Eq. (5), substituting in Eq. (7), and rearranging, also noting that

$$i_e(x) = \int_{x_0}^{x} j(x) C_e dx$$

$$\frac{d^2T}{dx^2} = \frac{C_e}{kA_e}\left[q_r(v, T) - q_0(x) - \frac{\rho_e C_e}{A_e}\left(\int_0^x j(v, T)\, dx \right)^2 \right] \tag{8}$$

Differential Eqs. (4) and (8) are coupled with the SIMCON (Ref. 1) tabular j, v, T (emitter) data (for fixed collector and cesium temperatures) including extrapolations to lower emitter temperatures than given in the basic SIMCON data. These relations are then solved with appropriate boundary conditions.

2. Problem Solution

Two second-order nonlinear ordinary differential equations are solved either as an initial value problem or as a boundary value problem depending on the knowledge of the initial voltage.

The basic equations (as derived above) are

$$\frac{d^2v}{dx^2} = -\frac{\rho_{\text{eff}} C_e}{A_e} j(v, T) \tag{9}$$

$$\frac{d^2T}{dx^2} = \frac{C_e}{kA_e}\left[q_r(v, T) - q_0(x) - \frac{\rho_e C_e}{A_e}\left(\int_0^x j(v, T)\, dx \right)^2 \right] \tag{10}$$

Tabular data, obtained from the SIMCON program, provide the thermionic relationships between j, v, T, and q_r that appear in the differential equations above. Given the values of v and T, the values of j and q_r are obtained by bivariate linear interpolation.

Three initial conditions are always specified, those being

$$T(x_0), \quad (dT/dx)_{x=x_0}, \quad \text{and} \quad (dv/dx)_{x=x_0}$$

If $v(x_0)$ is known, then the problem becomes simply an initial value problem. However, if $v(x_0)$ is not known, then boundary conditions at $x = x_f$ of either the form

$$(dT/dx)_{x=x_f} = -\frac{F C_e}{k A_e} \int_{x_0}^{x_f} q_0(x)\, dx \tag{11}$$

or

$$(dT/dx)_{x=x_f} = [T_s - T(x_f)]\, k_1/kA_e y \tag{12}$$

are specified.

In accordance with SPS 37-56, Vol. III, pp. 189–191, it should be noted that both ohmic heating and lead radiation are ignored in Eq. (12).

Normally, a boundary condition of the form (3) above is specified and a specific number of boundary value problems are solved for a particular range of the constant C. For each of these cases the following values are computed:

$$y = \frac{k_1}{Q_x}(T(x_f) - T_s) \tag{13}$$

$$v_{\text{net}} = v(x_f) - i(x_f)\, \rho_1 y \tag{14}$$

$$\rho = \frac{i(x_f)\, v_{\text{net}}}{C_e(x_f - x_0)} \tag{15}$$

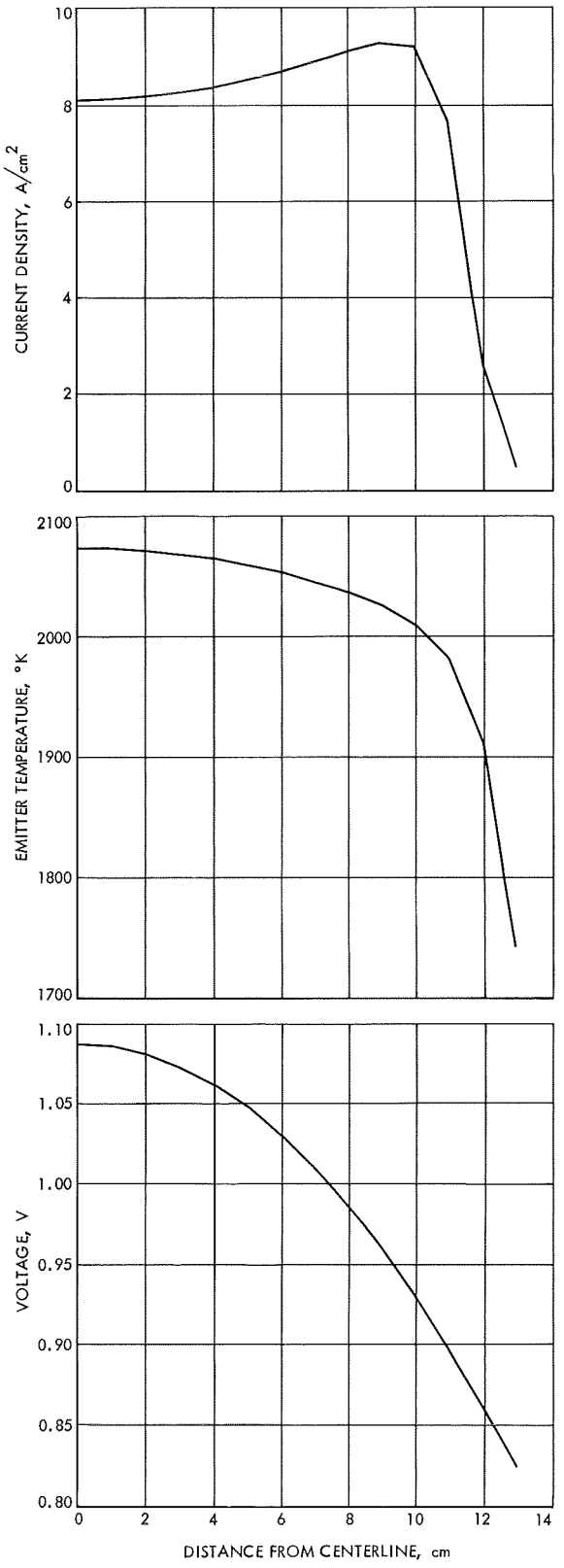


Fig. 7. Current density, temperature, and voltage distributions along diode length

and

$$\eta = \frac{p\,(x_f - x_0)}{\int_{x_0}^{x_f} q_0\,(x)\,dx} \tag{16}$$

It should be noted that by selecting a particular value of F, the heat out is specified, which then determines performance for a specified lead. By varying F over an appropriate range, performance of various leads is obtained. From these results one can then select the desired lead.

If a boundary condition of the type (4) is specified, the program computes the current density, voltage, and temperature profiles along the diode. Equations (13) through (16) are not evaluated.

If optimum lead parameters are desired, the user must provide as input two values for the constant F in Eq. (11). The program then solves a specified number of boundary value problems varying F (the lead heat loss fraction) over the range specified. The data generated then permits desired lead parameters to be determined.

Figure 7 shows typical program output for current density, temperature, and voltage distributions along the diode length. Figure 8 shows variation of net efficiency and power density versus average output current density.

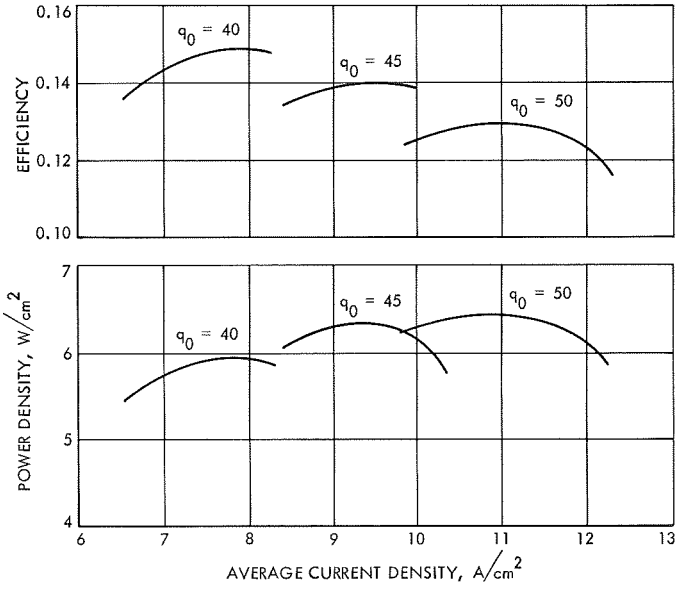


Fig. 8. Efficiency and power density vs average current density

Lead parameters are varied for each constant power input curve.

Reference

1. Wilkins, D. R., *SIMCON—A Digital Computer Program for Computing Thermionic Converter Performance Characteristics*, Report GESR-2109. General Electric Co., Pleasanton, Calif., Feb. 1, 1968.

C. Vaporizer Contamination by Mercury Impurities, *T. D. Masek and A. P. Rinder, Jr.*

1. Introduction

Mercury electron bombardment ion thrusters require the injected propellant to be in vapor form. Since the propellant is stored as a liquid, a vaporizer is required. Currently, most vaporizers are heated porous tungsten disks placed in the feed line. A typical assembly is shown in Fig. 9.

Proper operation depends upon the surface tension force at the mercury–tungsten interface. Since mercury does not wet tungsten ordinarily, the liquid can be contained on the upstream side of the disk to provide a stable surface for vaporization. If the liquid penetrates into or through the disk, the vaporization rate or flow rate cannot be easily controlled.

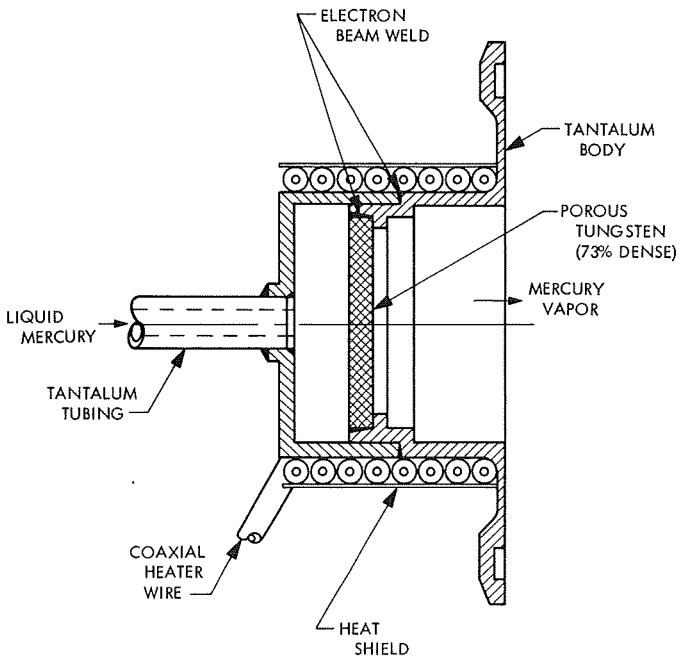


Fig. 9. Vaporizer assembly

Certain pressure constraints must be satisfied. The liquid pressure must exceed the vapor pressure at the liquid–vapor interface. This forces the liquid against the tungsten disk. However, the liquid pressure must not exceed that produced by surface tension at the disk or the liquid would be pushed into the disk. The surface tension pressure for present vaporizers is about 50 psi (8-μm average pore diameter).

2. Experimental Observations

During testing of a thruster and power conditioning unit combination, erratic operation was experienced. The arc, beam, and accelerator currents appeared to be noisy with irregular transients. During this type of operation the power conditioner was caused to "trip" or recycle, due to over-current conditions in one of these power supplies more frequently than normal. Observations indicated that the source of the current transients was the arc; i.e., the arc current reacted first causing the beam and accelerator to follow.

Further investigations showed that clamping off the propellant feed line, which was external to the vacuum chamber, eliminated the noise and transients within about one-half hour. This indicated that the vaporizer was allowing the passage of liquid mercury droplets. These would vaporize rapidly downstream of the vaporizer and cause bursts of pressure resulting in rapid increases in arc current. By clamping off the feed line, the liquid mercury pressure was reduced until the evaporation rate in the disk was equal to the liquid flow rate. When the pressure became too low, after 4 to 6 h, the vapor flow rate would decrease resulting in low arc, beam, and accelerator currents. This occurred when the liquid flow rate was lowered by the reduced liquid pressure. This allowed the vapor pressure to push the liquid–vapor interface to a point of lower temperature.

Two thrusters were used in the testing with similar vaporizer problems. Difficulty was encountered with both after about 24 h. However, after observing the problem, the thrusters remained in operation for 125 and 295 h, respectively. Upon disassembly of the vaporizers, deposits were found on the downstream surface of the tungsten. Photographs of the surfaces are shown in Figs. 10 and 11. Figure 10 clearly shows the deposit while Fig. 11 shows the deposit covered with liquid mercury. A section through the disk (Fig. 12) shows the penetration of the deposit.

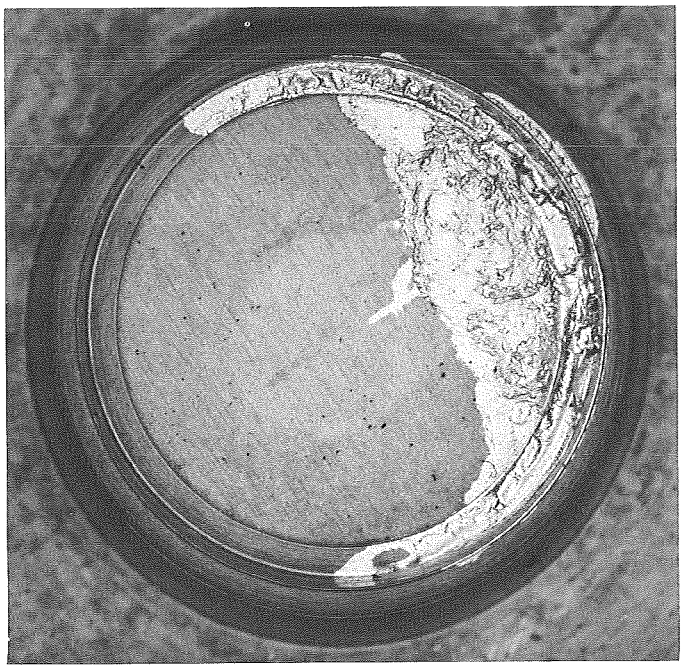


Fig. 10. Downstream face of thruster 1 vaporizer

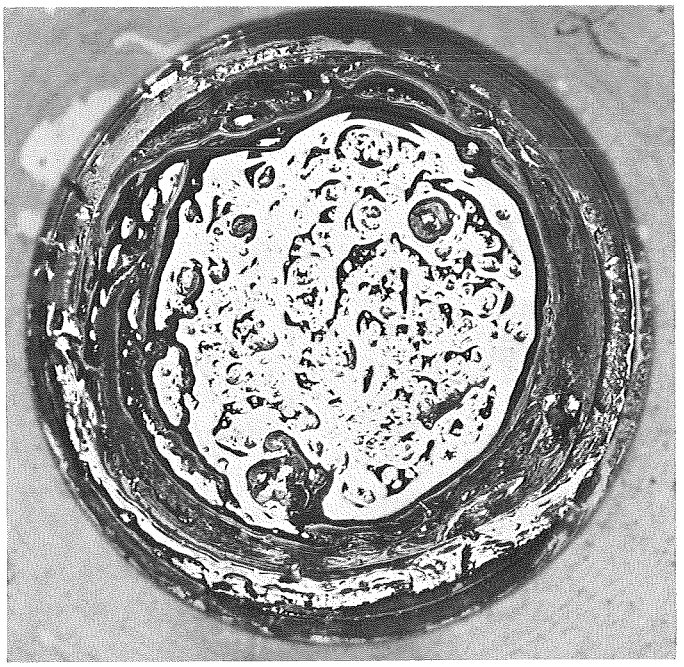


Fig. 11. Downstream face of thruster 2 vaporizer

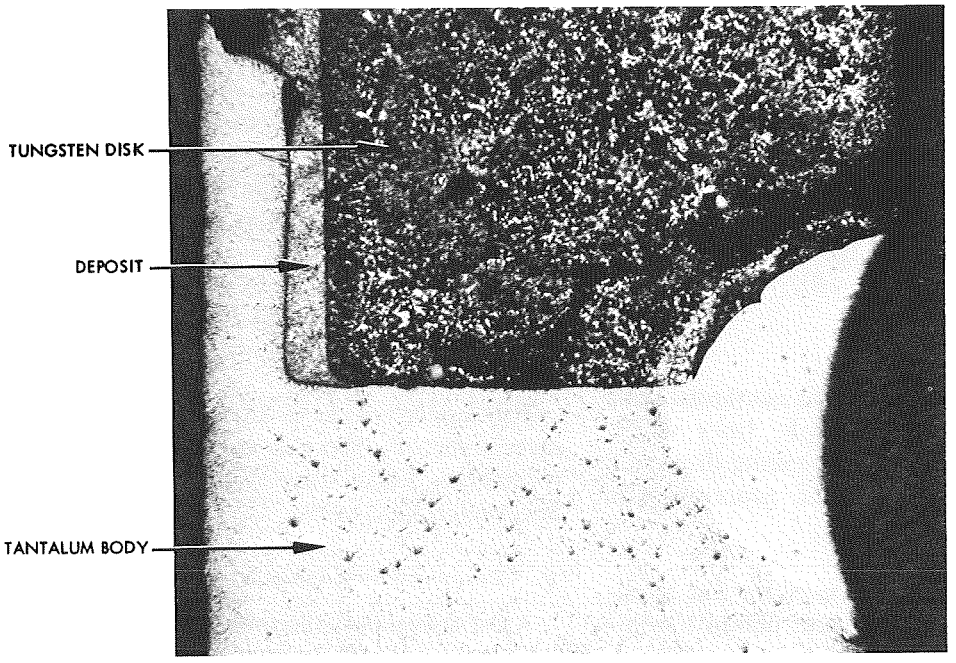


Fig. 12. Sectioned view of vaporizer (50 ×)

3. Analyses

The deposit was analyzed spectrographically with the following results:

Element	Deposit composition, %
Silver	67.4
Gold	0.086
Silicon	0.46
Mercury	32.0
Magnesium	0.0030
Aluminum	0.028
Tin	Trace (<0.001)
Palladium	0.0076
Copper	0.00049
Nickel	0.0024
Calcium	0.00066
Chromium	0.011
Other	Nil

Since an amount of silver as large as that shown (a few milligrams) could not have been present in the feed system, the mercury was analyzed.

Mercury samples were taken from the vaporizers after testing. In addition, samples of mercury from the laboratory supply (used in testing), JPL stores, and JPL chemistry section (distilled in a glass still) were obtained. Five samples were analyzed with the results shown in Table 3. It is apparent from these results that the silver and gold content was relatively high for the mercury samples used in testing (samples 1 to 3).

4. Conclusions

After the mercury analysis, testing was continued with new vaporizers and a mercury supply from which sample 5 of Table 3 was taken. After operating thruster 1 for about 900 h, no difficulty with the vaporizer was observed.

Based on this experience, mercury used for this application should be distilled thoroughly to obtain a composition with as few metal impurities as possible. A composition such as sample 5 can be obtained easily through proper distilling procedures. Specifically, the mercury should be washed in acid and distilled in a glass system.

Table 3. Composition of mercury samples

Element	Composition of sample,[a] parts/10^6				
	1	2	3	4	5
Silver	7100	9.5	14.0	0.54	Nil
Gold	15.0	0.37	0.084	Nil	Nil
Copper	0.039	Nil	0.059	0.049	0.20
Manganese	0.37	↓	Nil	Nil	Nil
Magnesium	0.056	↓	0.039	↓	0.048
Silicon	1.1	↓	Nil	↓	0.94
Iron	4.9	↓	↓	↓	Nil
Calcium	0.15	↓	↓	↓	↓
Chromium	2.3	↓	↓	↓	↓
Palladium	1.1	↓	↓	↓	↓
Nickel	1.2	↓	↓	↓	↓
Aluminum	0.21	↓	↓	↓	↓
Boron	Trace	↓	↓	↓	↓
Other	Nil	Nil	Nil	Nil	Nil
Mercury	← Remainder →				

[a]1 Thruster 1 (liquid in line at vaporizer).
2 Thruster 2 (liquid in line at vaporizer).
3 Laboratory bottle used in tests.
4 JPL stores 1-lb bottle.
5 JPL distilled using glass still.

D. Liquid–Metal MHD Power Conversion,

D. Bogdanoff and L. G. Hays

1. Introduction

Liquid–metal magnetohydrodynamic (MHD) power conversion is being investigated as a power source for nuclear-electric propulsion. A liquid–metal MHD system has no moving mechanical parts and operates at heat-source temperatures between 1200 and 1400°K. Thus, the system has the potential of high reliability and long lifetime using readily available containment materials such as Nb–1%Zr.

Fabrication of a 50-kWe NaK–nitrogen conversion system and of a 100-kWt cesium–lithium loop is nearing completion, and a new type of separator with no wall friction has been tested.

2. Cesium–Lithium Erosion Loop

A test system to evaluate high-velocity, high-temperature lithium erosion and to measure the thermodynamic conditions and nozzle performance of cesium–lithium mixtures is nearing completion. This system, which has previously been described (SPS 37-48, Vol. III, pp. 125–129), will operate initially at a maximum lithium temperature of

1250°K with a lithium velocity of 130 m/s. Lithium and cesium flow circuits and components are constructed of Nb–1%Zr alloy while ancillary lines and NaK heat rejection system are constructed of stainless steel.

All Nb–1% Zr weldments have been completed except for the final joining to the ducts of the electromagnetic pump. This achievement was highly significant since approximately 150 weldments were required, all of which had to be performed in a chamber with a purified argon atmosphere. The completed assembly is shown in Fig. 13, mounted on the support frame before application of thermocouples and radiation shielding.

A test of the heater power supplies, electrical bus, and water-cooled feedthroughs was completed with satisfactory results. At a current of 1900 A, which is a factor of 2 greater than the design value, the maximum temperature of this system was only 40°C above ambient. Thus, all subsystems (control, data readout, argon, vacuum, NaK heat rejection loop, heater power input) have been tested and are operable. The major tasks remaining before the loop is ready to fill are: mating all subsystems with the completed cesium–lithium loop, installation of thermocouples, load cell, and radiation shielding, calibration of instrumentation, and final checkout of the integrated system.

3. Nak–Nitrogen Conversion System

Empty-channel electrical tests of the generator have been completed, and final assembly of the conversion system is in progress.

4. Impinging Nozzle Experiment

The purpose of this experiment is to obtain a separator design using the impinging nozzle technique, with substantially reduced friction losses compared to those of the surface separators usually used in MHD power generation systems. If this objective is attained, it should be possible to obtain a 1 to 3 percentage point increase in cycle efficiency by employing the resulting impinging separator design.

a. General description. The apparatus used in the impinging nozzle experiment is shown in Fig. 14. Water and nitrogen, from a sump and a high-pressure supply, respectively, are metered using turbine-type flow meters and then fed to injectors at the upstream ends of the nozzles. The two-phase jets leave the nozzle exits and impinge on each other in the straight channel. The knife edges seen at the end of the straight channel collect the bulk of the water flow which has been concentrated towards the center of the channel, and the resulting jet is then deflected downward into the sump. The reaction of the jet on the test rig is measured using a load cell. The bulk of the gas flow (carrying, under some conditions, a substantial fraction of the water flow) is deflected outside the knife edge slot. Part of the water carried in this gas flow is caught in the secondary capture slots and fed to a weigh tank to allow the water mass flow to be determined. Most of the gas flow, still carrying considerable water, leaves the channel through the gas exhaust ports and is passed through a separator and discharged to the atmosphere. The water drain in the separator can be closed off to allow the separator water flow rate to be determined.

b. Parameters varied. The following parameters were varied in the experiments:

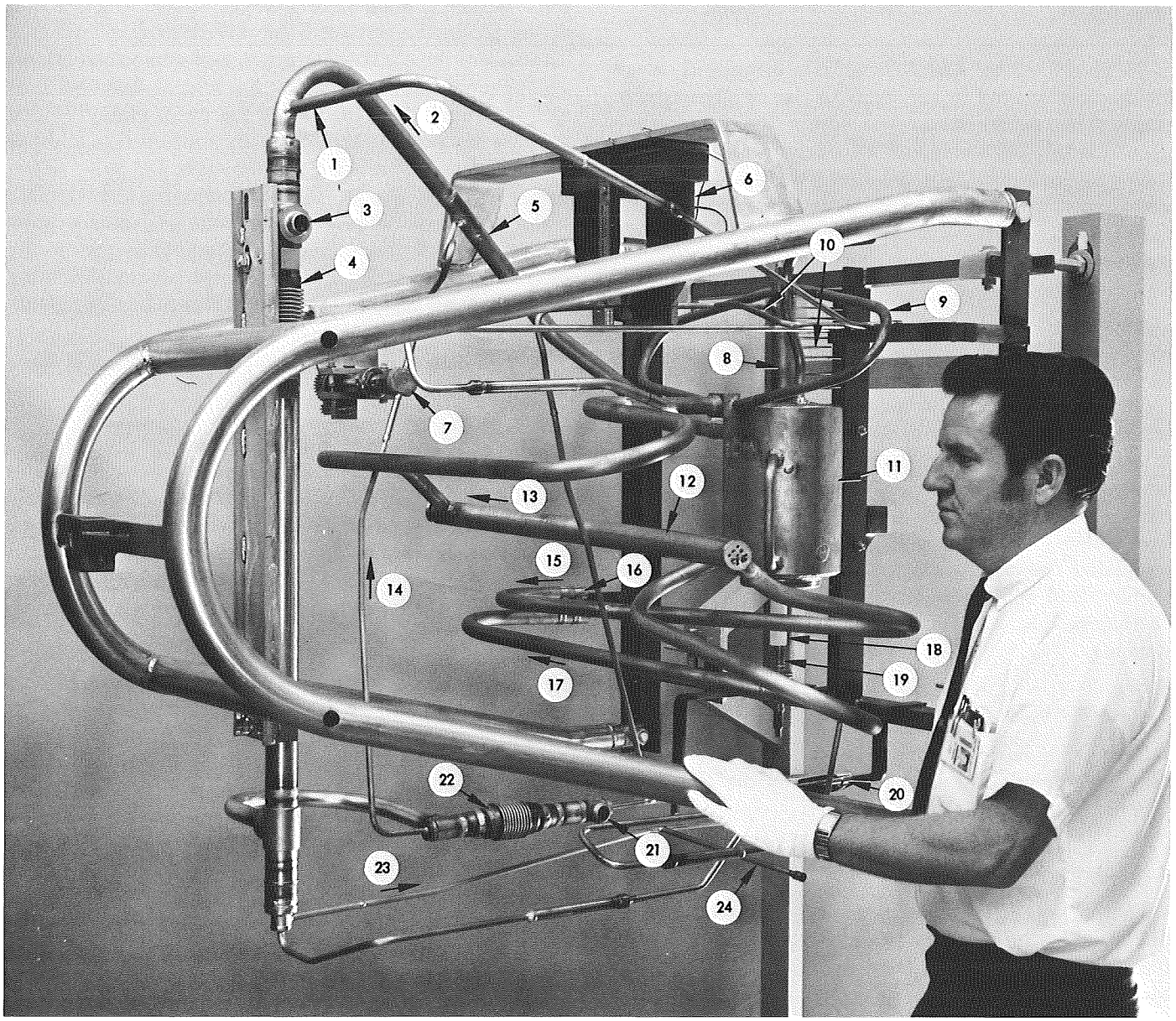
(1) The angle between the centerline of a nozzle and the centerline of the complete apparatus. Angles tested were 5, 10, 15, 20, and 30 deg. As the nozzle angle was varied, the length of the straight channel was changed to allow the knife edges to be located in the neighborhood of the geometric impingement point, the latter being located as shown in Fig. 15.

(2) The gap between the knife edges. Several sets of knife edge blocks were used to investigate gaps of 0.5, 0.75, 1.0, and 1.5 in.

(3) The distance of the knife edges downstream of the nozzle exits. This was varied in steps of 1 to 2 in. over 6 in. in the neighborhood of the geometric impingement point.

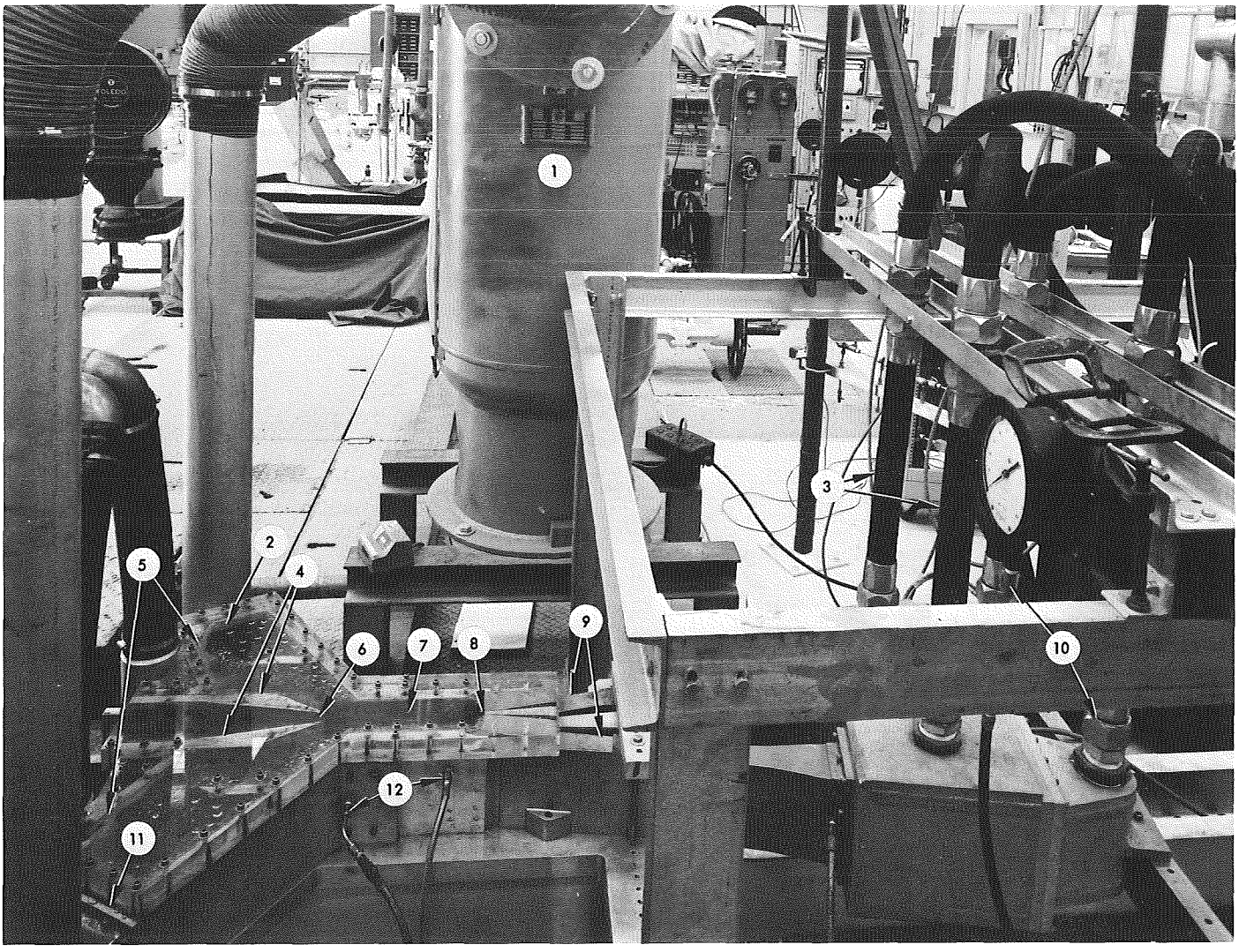
(4) The ratio of the nitrogen mass flow to the water mass flow. In the experiments to be described, the nitrogen and water mass flows were varied from 2 to 6 lbm/s and 80 to 170 lbm/s, respectively, giving gas/liquid mass flow ratios of 0.07, 0.04, 0.019, and 0.012.

(5) A limited series of experiments was conducted, turning the nozzles so that the large dimension (6 in.) was horizontal rather than vertical. This led to a separator channel 1.6 in. high × 13 in. wide to be contrasted with the original channel dimensions of 6 in. high × 4 in. wide. These tests were taken at one nozzle angle (15 deg) and consisted of mass flow probe traverses only (there were no knife edge blocks in the channel) at one distance downstream of the nozzle exit.



1 PRESSURE TAPS
2 Cs VAPOR
3 NaK OUTLET PORT
4 Cs CONDENSER
5 Cs DE-SUPER HEATER
6 Cs FLOW METER
7 Cs BYPASS VALVE
8 Li-Cs NOZZLE
9 Li NOZZLE INLET LINE
10 PRESSURE TAPS
11 Li-Cs SEPARATOR
12 Li HEATER
13 Li LIQUID
14 Cs LIQUID
15 Li LIQUID
16 Li PUMP LINES
17 Li LIQUID
18 LEVEL INDICATOR
19 THRUST EXTENSION
20 Cs PUMP LINES
21 NaK INLET PORT
22 Cs SUBCOOLER
23 Cs LIQUID
24 PRESSURE TAPS

Fig. 13. Cesium–lithium erosion loop



1 WATER SEPARATOR
2 GAS DISCHARGE PORT
3 NITROGEN FEED LINES
4 KNIFE EDGE BLOCKS
5 SECONDARY CAPTURE SLOT
6 PRIMARY CAPTURE SLOT
7 SEPARATOR CHANNEL
8 NOZZLE EXITS
9 NOZZLES
10 WATER SUPPLY LINES
11 GAS DISCHARGE PORT
12 STATIC PRESSURE TAPS

Fig. 14. Impinging nozzle test apparatus

c. Measurements using knife edge blocks. Figure 16 shows the fraction of the water flow bypassing the primary capture slot (R_B) as a function of the distance of the knife edges from the nozzle exit x (Fig. 15) for four nitrogen-to-water flow ratios R_F at nozzle angle = 20 deg, and knife edge gap = 0.75 in. As the relative nitrogen flow ratio (R_F) is decreased, Fig. 16 shows generally improving separator performance (decreasing R_B values). This same dependence was observed for all nozzle angles and knife edge gaps tested, except for the smallest knife edge gap (0.5 in.), where the R_F = 0.012 data generally showed higher bypass ratios than the R_F = 0.019 data. A possible explanation for this dependence is as follows: At the lower R_F values, the velocities and accelerations of the gas flow in the nozzles are lower and hence the absolute velocity difference between the water droplets and the gas tends to be less. Since the surface tension forces at larger droplet diameters are sufficient to withstand the lower aerodynamic forces, lower R_F values would be expected to produce larger water droplets. It is to be noted here that while the absolute velocity differences are lower for lower R_F values, the fractional velocity differences (velocity

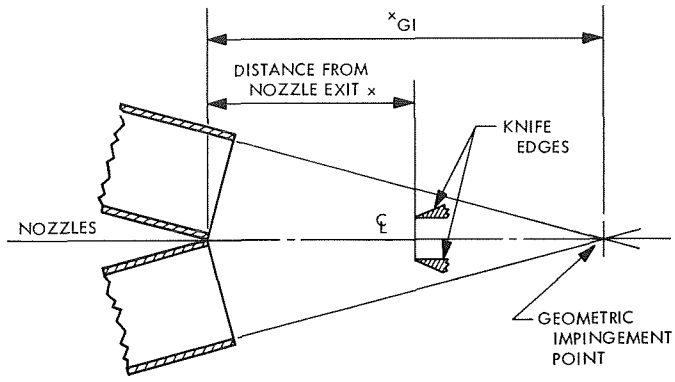


Fig. 15. Definition of geometric impingement point and distance from nozzle exit

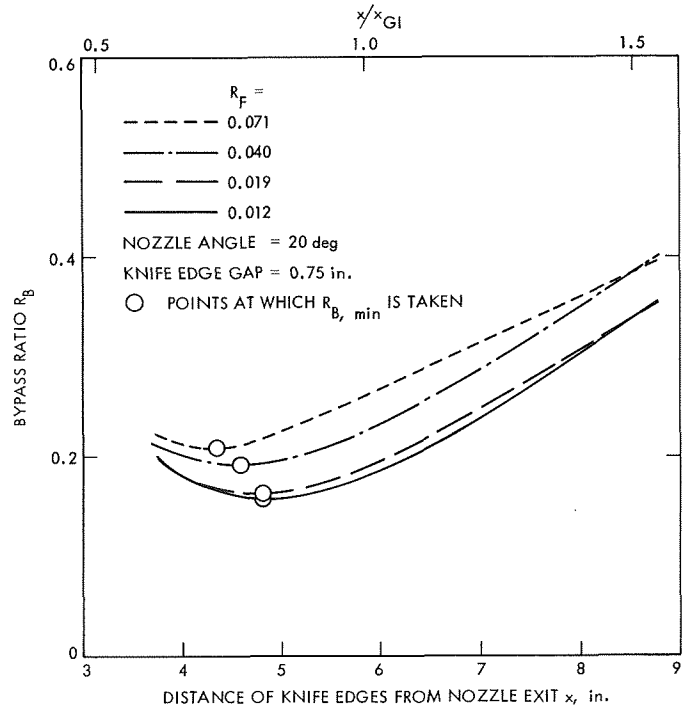


Fig. 16. Bypass ratio vs distance of knife edges from nozzle exit

difference divided by gas velocity) are larger. The flow separator functions by deflecting the gas flow outwards just beyond the nozzle exits, while the water droplets from the two nozzles continue on their converging paths and coalesce in the center of the channel. Thus, at lower R_F values, the larger water droplets should be less affected by the deflecting gas stream, giving better separator performance. One obvious limitation of the above discussion is that droplet collisions are obviously very important in the real flow (typical mean free path 0.01 in.).

The variations of R_B with R_F shown in Fig. 16 (up to 30%) are typical for all nozzle angles and knife edge gaps except for the largest (1.5 in.) gap, where the variations of R_B with R_F may be up to 50%. Since R_B varies typically <30%, while R_F varies over a factor of 6, the local mass flow ratios at various points in the separator channel (particularly towards the center of the channel) must tend to be proportional to the overall mass flow ratios. In particular, the maximum concentration of water in the center of the separator channel appears to be roughly proportional to the water concentration at the nozzle exit, rather than to approach a fixed maximum value for all R_F values.

From Fig. 16, the lowest values of R_B ($R_{B,\min}$) are seen to occur at about 0.8 of the distance of the geometrical impingement point from the nozzle exit, i.e., at $x/x_{GI} \sim 0.8$. Denoting x/x_{GI} at the minimum of R_B by Z_M, Z_M tends to increase with increasing nozzle angles and knife edge gaps, rising from 0.65 at an angle of 10 deg and a gap of 0.5 in. to 0.85 at an angle of 30 deg and a gap of 1.0 in. At the largest gap (1.5 in.), Z_M is substantially larger, ranging from 0.9 at an angle of 10 deg to 1.15 at an angle of 30 deg. Thus, there is a region, typically from $x/x_{GI} = 0.75$ to 1.05, where the water flow in the most central regions of the channel has already reached its maximum concentration and is starting to spread out again, while in the edge of the flow the water is still moving towards the center.

For each flow ratio (R_F) and knife edge gap, plots of $R_{B,\min}$ vs nozzle angle were made to determine the optimum nozzle angle. In all cases, a nozzle angle of 20 deg yielded the lowest values of $R_{B,\min}$.

d. Measurements using mass flow probes. A water mass flow probe was employed which consisted of a sharp-edged tube of 0.040-in. diam connected to a 0.25-in.-diam support tube and discharging into a graduated cylinder. The probe was traversed across the separator channel. For these measurements the knife edge blocks were either removed from the channel or were located well downstream of the probe. In general, R_B values computed from the mass flow profiles were lower than those measured using the secondary capture slot and separator water mass flows with the knife edges in place. The differences ranged from 5% at the larger gaps to 15% at the smaller gaps. Two factors possibly contributing to these differences were: (1) the diversion of the gas flow by the outer surfaces of the knife edge blocks may carry some of the water away from the knife edge slot; and (2) with the knife edges in place, the

pressure in the separator channel is raised considerably (as much as 7 psi) above atmospheric. Both of these effects would not occur when the mass flow profiles are taken. Possibly, by building the separator capture slot out of two thin plates instead of the thick knife edge blocks presently used, the R_B values of the separator could be made to approach those values calculated from the mass flow probe data. Thus, further improvement of R_B by 5–15% may be possible.

The mass flow probe profiles taken at $x = 8.7$ in. confirm the data obtained using the knife edge blocks with respect to the lower R_B values at lower relative nitrogen flow rates (lower R_F values). Figure 17 shows profiles of probe water mass flow M_P for a nozzle angle of 10 deg, $R_F = 0.019$, and with the probe located 5.7, 7.2, 8.7, and 11.0 in. downstream of the nozzle exit (note that the maximum water density is slightly off the channel centerline). It can be seen that the concentration of water towards the center of the channel continues to increase between $x = 5.7$ and $x = 8.7$ in., i.e., M_P increases in the central regions of the flow and decreases in the edges. Between $x = 8.7$ and 11.0 in., M_P continues to decrease in the edges; however, the concentration of water in the central regions of the flow is decreasing from its highest values. Over the center 0.5 in. of the flow, the greatest concentration of water occurs at $x = 8.7$ in. and $x/x_{GI} = 0.62$. These characteristics of the water flow distribution in the channel, as measured using the mass flow probe, are in close agreement with those obtained using the knife edge blocks.

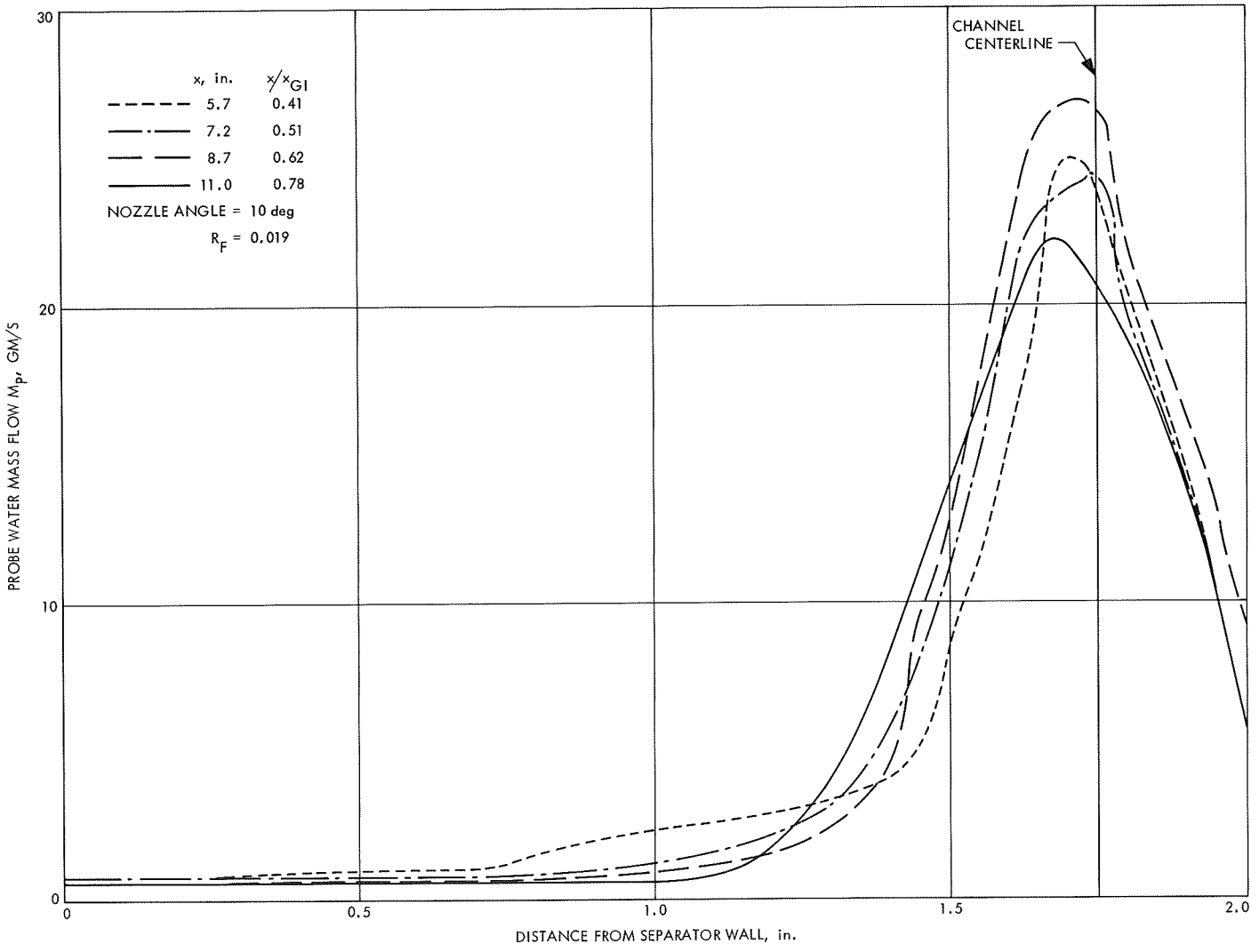


Fig. 17. Profiles across separator channel of probe water flow rate

e. Data taken with the large dimension of the nozzles horizontal. The data with the large dimension of nozzle horizontal consisted only of mass flow probe profiles taken with a nozzle angle of 15 deg, $R_F = 0.019$, and the probe 23 in. downstream of the nozzle exit ($x/x_{GI} = 1.00$). Comparisons between these profiles and those taken with the 6-in. nozzle dimension vertical, though somewhat uncertain, show that the degree of concentration of the water obtained with the 6-in. nozzle dimension horizontal and at $x/x_{GI} = 1.0$ is roughly equivalent to that obtainable with the 6-in. nozzle dimension vertical if the same R_F values (0.019) and nozzle angles (15 deg) were used for both cases. This is somewhat surprising, in view of the fact that calculations for droplet trajectories in suddenly deflected gas streams indicate that droplets with diameters in the 0.002-in. range (which should be fairly plentiful, since calculations based on a limiting Weber number for droplet breakup of 6 indicate that the largest droplets which can survive the high accelerations in the nozzle throat have diameters of ~0.005 in.) should not be able to move transversely to the gas flow more than about 2 in. Two possible effects may contribute to the relatively high performance of the separator with the 6-in. nozzle dimension horizontal, as compared to what might be expected on the basis of the trajectory calculations. First, visually, there are some indications that there is a large gas vortex near the channel wall. Such a vortex could produce centrifugal forces which are more effective in moving the droplets transverse to the channel centerline than the turning of the gas flow alone. Secondly, as the smaller droplets start to turn and follow the gas stream, they may collide and coalesce with larger droplets which are continuing on their straight paths towards the channel centerline.

f. Conclusion. The impinging-jet separator provides a three- to four-fold concentration increase of the liquid without friction loss. Cycle calculations show that a 1 to 3 percentage point increase in cesium–lithium cycle efficiency could be obtained using the impinging-jet separator instead of a surface impingement separator. The exact increase would depend on the losses in handling the higher vapor volume in the collected flow and greater liquid containment in the exhaust vapor with the impinging-jet separator.

E. Study of Cesium Reservoir Control for Thermionic Reactors, *R. Hermsen and H. G. Gronroos*

1. Introduction

This article reports the results of a conceptual study of cesium reservoir temperature control in thermionic reactor systems. This problem is of interest because of the sensitivity of the thermionic conversion efficiency to only small changes in reservoir temperature, and because of the possibility of unstable operating modes if the temperature is too low. These investigations are part of continuing analytical and experimental studies of thermionic reactor physics, reactor design, and applications for spacecraft nuclear power. Articles specifically dealing with the dynamic aspects of a thermionic reactor powerplant are published in SPS 37-45, Vol. IV, pp. 151–162, SPS 37-55, Vol. III, pp. 220–234, and Refs. 1 and 2.[1]

2. General Considerations

Analytical and experimental work has shown that a high-performance thermionic converter operating at peak efficiency in the emitter temperature and current density ranges normally used can become unstable under an increasing heat input disturbance (SPS 37-57, Vol. III, pp. 173–177, and Refs. 2 and 3). This phenomenon, called thermionic burnout, is illustrated with reference to Fig. 18. The figure shows the current density J and corresponding emitter heat flux Q_E as a function of emitter temperature T_E with the cesium reservoir temperature T_R as parameter. The curves are for a fixed diode voltage V_D at 0.6 V and a fixed collector temperature T_C at 1000°K. This anticipates a thermionic reactor application with constant output voltage control and with all converters operating in the same state. If another control voltage had been selected, similar curves with a maximum point would have been obtained. This also holds if a constant load line had been used instead of a constant voltage line. The curves in Fig. 18 are based on data obtained from the SIMCON code for a diode with a tungsten emitter, a molybdenum collector, and a 0.010-in. (0.025-cm) interelectrode gap (Ref. 4).

Point A on the $T_R = 610$°K curve in Fig. 18 is a maximum efficiency point with heat input of 72 W/cm². Only a 4% increase in heat flux (caused by a reactivity perturbation in the reactor) is sufficient to increase the emitter temperature past the maximum point B at 75 W/cm². At this point the cesium begins to desorb from the emitter surface. This in turn lowers the current density, hence less electron cooling which leads to a further buildup in emitter temperature and so on. The increase in emitter temperature would continue until

[1]Also Shapiro, J. L., and Gronroos, H. G., "Stability Analysis of a Thermionic Reactor," paper to be presented at the 1969 Winter Meeting of the American Nuclear Society, San Francisco, Calif., Nov. 30–Dec. 4, 1969.

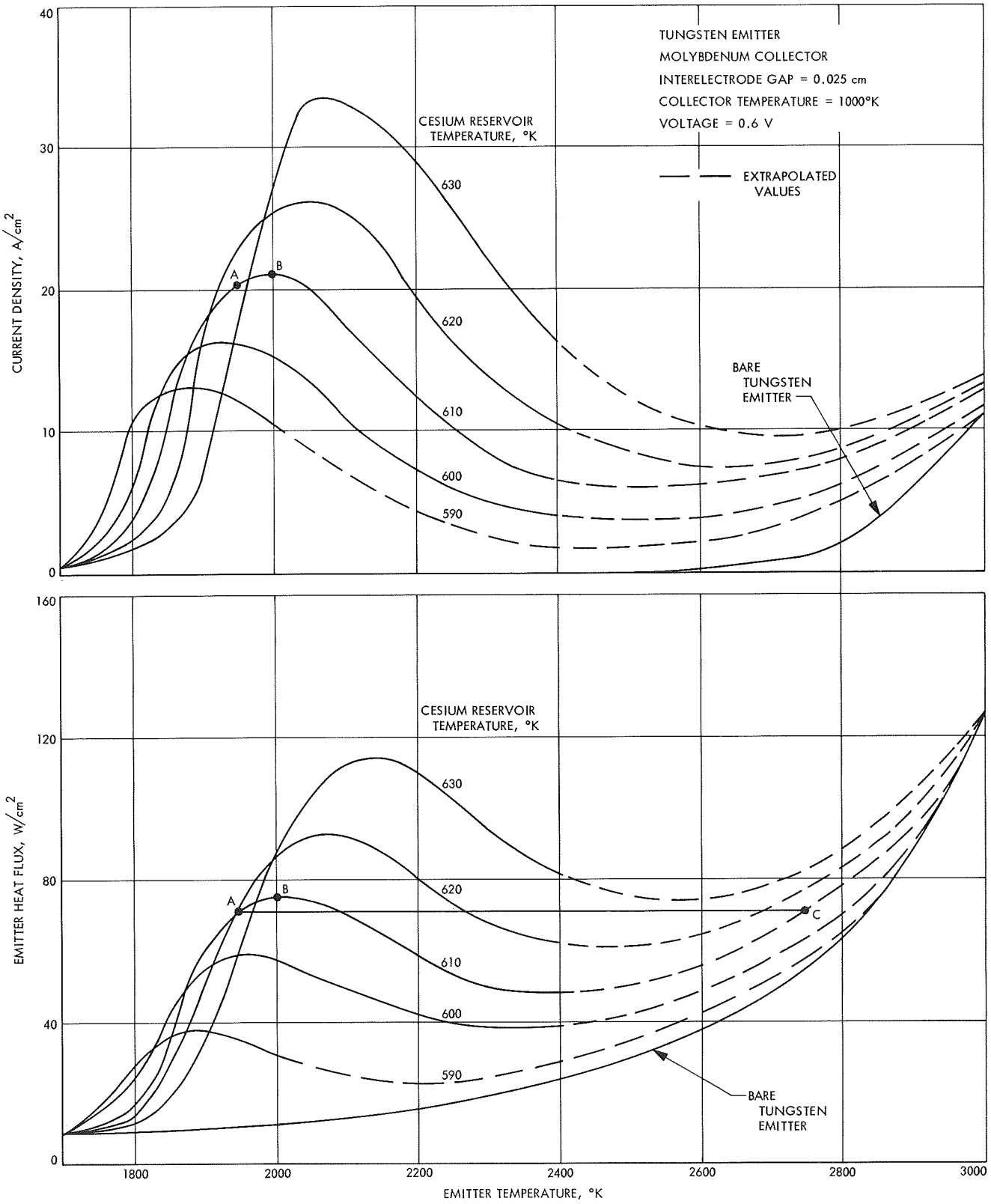


Fig. 18. Current density and emitter heat flux as a function of emitter temperature

point C is reached if there were no negative temperature coefficient of reactivity causing the reactor to decrease its power as the temperature increases. Since this is generally the case, the reactivity coefficients, together with the thermal lags in the system, can cause the system to oscillate around the maximum point. Obviously a drop in cesium reservoir temperature of a few degrees would have the same effect.

The cesium desorption point occurs at higher emitter temperatures and heat fluxes as the cesium reservoir temperature increases. If no reservoir temperature control were employed, this temperature would have to be set at higher than optimum at a noticeable loss in conversion efficiency (1 to 3 points) to provide an adequate margin of safety. Thus, it is desirable to design a reservoir controller that allows operation near maximum efficiency for all diode states in the operating range, and that prevents the possibility of burnout.

Two control schemes have been investigated. For one scheme current, voltage, and reservoir temperature were used as states that define controller operation; in the other scheme the emitter heat flux was used instead of the voltage. Variations in the heat flux are proportional to neutron density changes, which in turn are controlled by the reactivity. The variables most useful for defining the diode state are current and emitter temperature; however, the latter variable is not readily measurable.

For actual thermionic diodes the maxima for the heat flux are not as pronounced as indicated by Fig. 18 due to additional losses and non-uniform emitter temperature distribution not considered in the SIMCON code calculations. The SIMCON data are used for illustrative purposes to determine the feasibility of various cesium reservoir control schemes.

3. Reservoir Controller Employing Current and Voltage States

The cesium desorption points (T_{RS}) have been plotted on a reservoir temperature versus current density graph for different diode voltages as shown in Fig. 19. Superimposed are also constant emitter temperature lines. The stable region is above the parametric curves. The stability boundary for the constant voltage curves may be approximated by the following linear relationship:

$$T_{RS} = M_1 J + M_2 V_D + B_1 = 2.15 J + 67.5 V_D + 528 \quad (1)$$

Figure 20 shows the mechanization of Eq. (1) for a system with a proportional controller and a lag element describing the dynamics of the reservoir. The stability boundary is relatively unaffected by collector temperature variations between 850 to 1150°K, and the linearization over the ranges shown in Fig. 19 sets the cesium desorption point within ±4°K accuracy.

The disadvantages of controlling the system on current and voltage measurements are twofold. Firstly, the voltage, although precisely measurable, can vary considerably from diode to diode and from calculated values. Secondly, if the diode is operating in the burnout region, the reservoir control loop becomes inoperative since then both current and voltage decrease with increasing emitter temperature.

4. Reservoir Controller Employing Current and Heat Input States

The physics of the thermionic conversion phenomena imply that knowledge of the emitter temperature and electrical current would define the appropriate cesium reservoir temperature over a reasonable range about the nominal operating point (i.e., for a given diode design and collector temperature). Unfortunately, as mentioned above, the emitter temperature is not readily measured, but measurement of the heat input to emitter and the current density can be used to infer the steady-state emitter temperature. In a nuclear reactor application this implies measurement of the neutron flux (the change in coolant temperature drop across the reactor is undesirably slow). A linear expression approximating the relationship of the emitter temperature T_E to the emitter heat flux Q_E and the current is

$$T_E = M_3 Q_E + M_4 J + B_2 = 16.4 Q_E - 48 J + 1660 \quad (2)$$

The approximating linear relationship for T_E as a function of the reservoir temperature stability boundary and J is obtained from Fig. 19. The result is

$$T_{ES} = M_5 J + M_6 T_R + B_3 = -7.7 J + 9.4 T_R + 3560 \quad (3)$$

The difference between the inferred emitter temperature, Eq. (2), and the emitter temperature at the cesium desorption point, Eq. (3), is used to control the reservoir temperature, i.e.,

$$\Delta T_E = e = 16.4 Q_E - 40 J - 9.4 T_R + 5220 \quad (4)$$

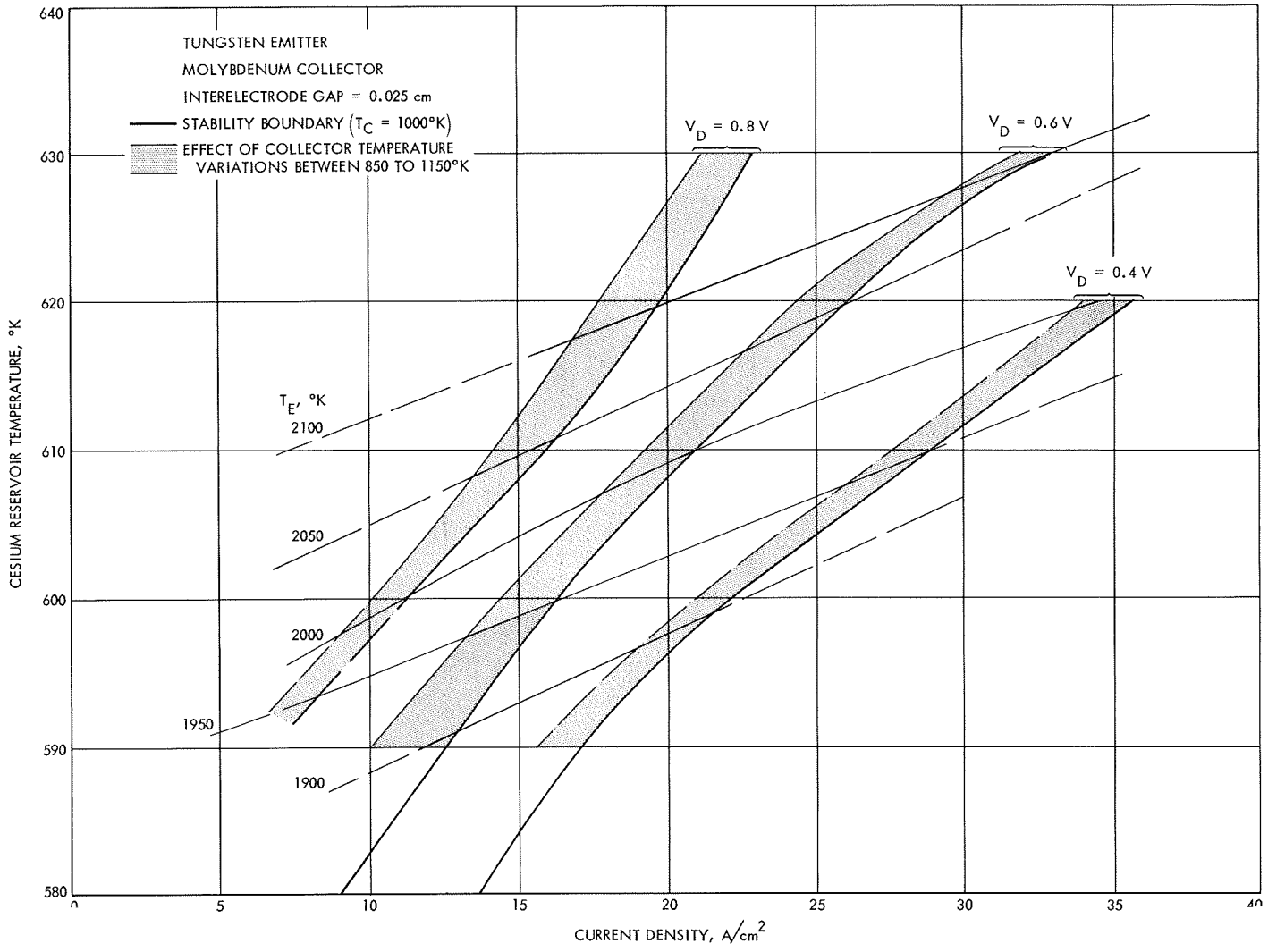


Fig. 19. Reservoir temperature at cesium desorption vs current density

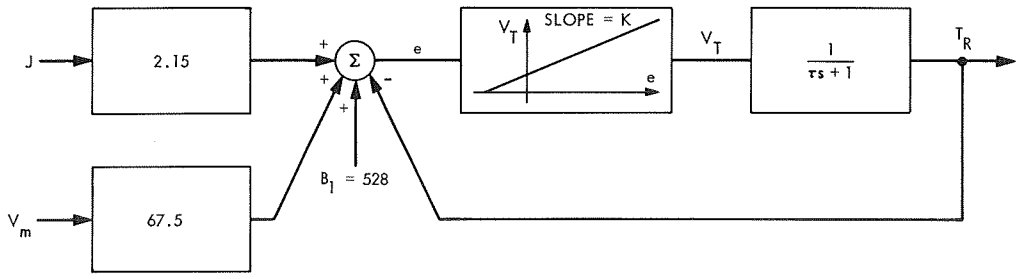


Fig. 20. Block diagram of cesium reservoir controller employing current and voltage states

This linear relationship controls T_R to within 5°K of the SIMCON data. The resulting control system is shown in Fig. 21.

It is of interest to compare Eqs. (1) and (4). In the stable steady-state operating region, both equations are essentially equivalent. Equation (1) controls the cesium reservoir temperature by measuring the diode outputs J and V_D, while Eq. (4) controls the reservoir temperature by measuring a diode input variable, Q_E, and an output variable, J. The negative current coefficient in Eq. (4) conforms more closely to the basic nature of the thermionic burnout instability. A decrease in current implies less electron cooling and hence a higher emitter temperature, which in turn requires a higher reservoir temperature. As a consequence, the system of Fig. 21 performs satisfactorily in the presence of disturbances which shift the operating point to the unstable region. Since the diode heat input changes instantaneously with neutron flux, and since the emitter temperature follows with a time lag, the system anticipates diode operation changes due to reactivity disturbances.

The setting of the gain constant K and requirements on the cesium reservoir time constant τ (Fig. 21) depend on the specific reactor design. The bias should be set so that there is some margin to the stability boundary, say, 10°K higher than T_{RS}. The reservoir dynamics, as shown in Fig. 21, has the transfer function $K/(\tau s + 1)$ which is a type 0 system. Therefore, there is a steady-state error for a step change in J and Q of $1/(1 + 9.4K)$. Considering a 40% change in load for the plant model studied in Ref. 1 and a reactivity rate constraint of $\dot{\rho} \leq 10\ \phi$/s, then $(dT_E/dt)_{max}$ is ≤ 7°K/s for 5 s. These conditions require the reservoir temperature change to be 1°K/s. For a gain constant of $K = 1.6$ and a time constant of $\tau = 1$, one obtains $(\Delta T_R/dt)_{max} = 1.4$°K/s. The response is not materially improved by further increasing K and decreasing τ.

5. Conclusions

This preliminary design study indicates that it is feasible to control the cesium reservoir temperature, and to maintain a high conversion efficiency over a range of diode operating states. The conversion efficiency is a strong function of the reservoir temperature and a thermal runaway condition exists. If no control is applied, the reservoir temperature must be set with a margin for worst-case operating conditions with an attendant loss in conversion efficiency. The final controller design depends on the specific application. The control accuracy depends mainly on the following: (1) thermionic performance data and diode performance reproducibility, (2) range of operating conditions, (3) diode matrix configuration and reactor power distribution, (4) errors in a linearized analysis, and (5) steady-state and dynamic errors of reservoir temperature control system. These error sources imply that the reservoir temperature should be set 10°K or more above design optimum and that little would be gained by a design taking nonlinearities into account.

Further investigation of the above results by analytical and experimental simulation studies is planned.

References

1. Weaver, L. E., et al., *A Control System Study for an In-Core Thermionic Reactor*, Technical Report 32-1355. Jet Propulsion Laboratory, Pasadena, Calif., Jan. 15, 1969.
2. Gronroos, H. G., and Davis, J. P., "Stability and Control Considerations for Thermionic Reactors," paper presented at the Second International Conference on Thermionic Electrical Power Generation, Stresa, Italy, May 27–31, 1968.

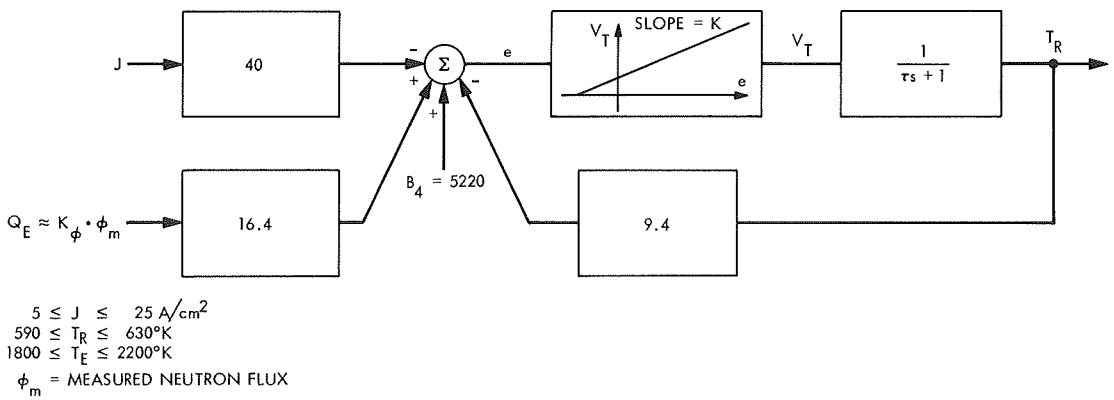


Fig. 21. Block diagram of cesium reservoir controller employing current and emitter heat input states

3. Schock, A., "Effects of Cesium Pressure on Thermionic Stability," paper presented at the Second International Conference on Thermionic Electrical Power Generation, Stresa, Italy, May 27–31, 1968. Also available in IEEE Conference Record of 1968 Thermionic Conversion Specialist Conference, Framingham, Mass., Oct. 21–23, 1968.

4. Wilkins, D. R., *SIMCON—A Digital Computer Program for Computing Thermionic Converter Performance Characteristics,* Report GESR-2109. General Electric Co., Pleasanton, Calif., Feb. 1, 1968.

F. Combined Theoretical and Preliminary Experimental Results of Heat Transfer to Anodes, *P. F. Massier and T. K. Bose*

1. Introduction

An experimental investigation is in progress to determine the heat transfer and fluid dynamics associated with electrical discharges that are influenced by transverse gas flow. This work was initiated because the heat transfer to some of the elements of plasma propulsion devices, particularly the anode, can be significantly higher than that encountered in chemical propulsion thrust chambers. Plasma temperatures may exceed 25,000°R, whereas common electrode materials, such as copper, melt at temperatures less than 2500°R. Furthermore, the impingement of electrons on the anode can produce an amount of heating which greatly exceeds that caused by convection from the directed motion of the plasma. Thus, in order to provide the correct amount of cooling for magnetoplasmadynamic arcs, magnetohydrodynamic generators and accelerators, and arc jets, it is essential to be able to predict the total heat transfer to the electrodes by taking into account the various contributing mechanisms.

This investigation is one part of a more comprehensive combined experimental and theoretical study of the various mechanisms and their interactions that affect the heat transfer from plasma flows. These effects include ionization, large variations in properties across the boundary layer, acceleration, nonadiabatic core flows, swirl, and applied magnetic and electric fields. A theoretical study of plasma boundary layer flow over electrodes, for example, is summarized in SPS 37-56, Vol. III, pp. 201–203.

The following discussion pertains to a theoretical analysis combined with preliminary experimental results of heat transfer to a matrix of flat, electrically and thermally insulated anode segments with an electric field applied to only one of the upstream segments. Thus, there was no current flow to the downstream segments. Figure 22 is a photograph of the experiment showing two streams of luminous plasma flow between the flat copper anode segments above and the rounded-tip tungsten cathodes below. There are 16 pairs of electrodes. The two luminous streams of plasma are actually a single electrical discharge between one pair of the electrode segments. This discharge has been deflected from a vertical plasma column into an elongated U-shape by the motion of argon gas from left to right; however, the view does not extend sufficiently downstream to show that the two plasma streams are actually connected. The gas flow entered the region between the electrodes from a convergent nozzle. As indicated on the photograph, the discharge was attached to the downstream edge of the anode. In this experiment each of the electrode segments was individually cooled and the heat transfer rate was determined from measurements of the coolant flow rate and temperature rise as the coolant flowed over the backside of the anode.

Furthermore, the total heat transferred by convection, diffusion, and radiation from the deflected plasma stream to other segments downstream, which did not have an applied electric field, could also be determined. Therefore, with these experiments some of the contributing factors to the total heat transferred to anode surfaces can be established, namely, electron bombardment and combined forced convection, diffusion, and radiation.

The conditions of the test shown in Fig. 22 were as follows: the pressure was 300 torr, the velocity of the argon gas was 50 ft/s, and the current was 39 A. The size of each anode segment is 0.25 × 0.25 in., the space between adjacent segments is 0.040 in., and the distance between the copper anodes and the tungsten cathodes is 1.0 in.

The experimental results to be discussed were acquired using argon at a nominal pressure of 50 torr, a current of approximately 40 A, and gas velocities between 55 and 257 ft/s.

The symbols used in this article are defined in Table 4.

2. Theoretical Predictions

The mathematical theory, which is used to compute the heat transfer to the anode, is based on the physical model that a very thin free-fall region exists adjacent to the wall within the plasma boundary layer, where the temperature of the heavy particles (atoms, ions, etc.) is equal to the

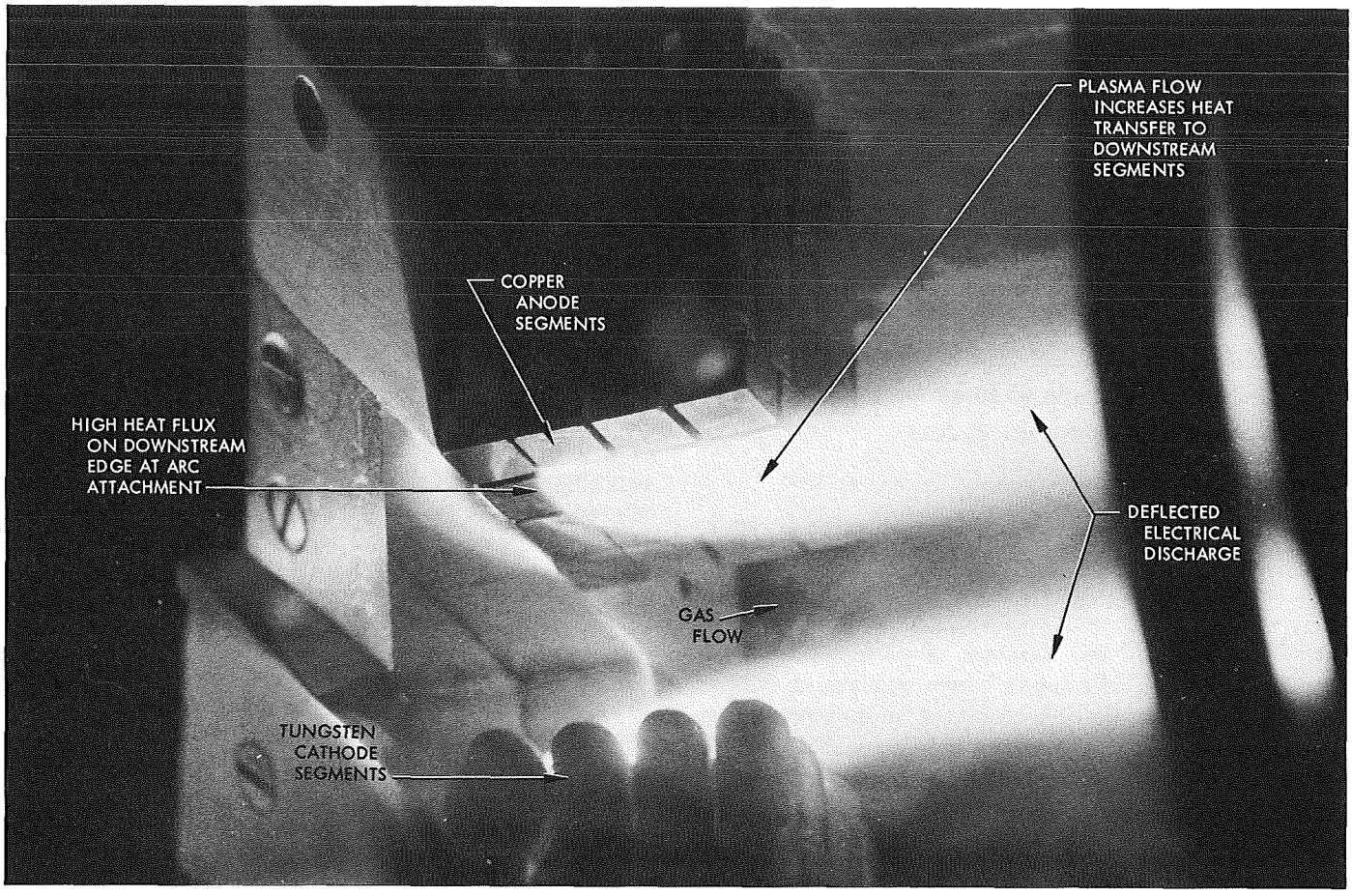


Fig. 22. Plasma flow over electrode surfaces

temperature on the wall surface and the gas velocity is zero. A continuum region outside the free-fall region is assumed in which the flow is laminar and steady. In addition, the electrons, falling freely through a sheath from one mean free path without any collisions, have at the free-fall edge, which is outside the sheath edge, a vanishingly small mass density and a temperature equal to that of the electrons in the free stream.

The general relation for the total heat flux at the wall, neglecting the heat transfer by radiation, as derived by Bose (Ref. 1), is

$$q = -k_h \nabla T - k_e \nabla T_e - \rho D_{\text{amb}} I_{mi} \nabla g_i - \frac{5}{2} \frac{R^*}{m_e} \rho D_{\text{amb}} T_e \nabla g_e + \frac{5}{2} \frac{k_B}{e} j T_{e_\infty} + j(\phi_s + \phi_w) \quad (1)$$

In Eq. (1) the effect of the second and fourth terms on the right side is found to be small in comparison to the other terms and hence can be neglected. Also, it can be seen that the heat flux relation has two parts, one of which is independent of the current and may be described as a combined convective and diffusive heat flux,

$$q_{c,d} = -k_h \nabla T - \rho D_{\text{amb}} I_{mi} \nabla g_i \quad (2)$$

The other is dependent on the electric current density and may be described as a heat flux resulting from the current,

$$q_j = j\left[\frac{5}{2} \frac{k_B}{e} T_{e_\infty} + \phi_s + \phi_w\right] \quad (3)$$

An expression for Nusselt number for flow over a flat plate can be obtained from Eq. (2) by nondimensionalizing and transforming from the natural x,y coordinates into the s, ξ coordinate system by using a transformation proposed

Table 4. Nomenclature

D_{amb}	ambi-polar diffusion coefficient
e	charge of an electron
g_j	species mass density ratio $= \rho_j/\rho$
I	current, A
I_{mi}	ionization potential per unit mass ion
j	current density, A/in.^{2}
k	thermal conductivity coefficient, Btu/s-in.-°R
k_B	Boltzmann constant
ℓ	segment number
L_x, L_y	anode dimensions, in.
m	mol mass, lb/mol
Nu	Nusselt number, Eq. (4)
N_1	dimensionless number, Eq. (5)
p	pressure, torr
q	heat flux, Btu/in.^{2}-s
Q	total heat load, Btu/s
$Q_{c,d}$	combined convective and diffusive heat load, Btu/s
Q_J	heat load caused by current, Btu/s
Re	Reynolds number, Eq. (5)
R^*	universal gas constant
Sc	Schmidt number, Eq. (5)
T	temperature, °R
U_∞	free-stream velocity, ft/s
x,y	coordinate directions parallel to surface along and normal to flow
x_0	distance from leading edge to origin of thermal boundary layer, in.
α	$\alpha\,(Pr)$
ρ	mass density, lb/ft^3
ϕ_s	anode potential with respect to plasma, V
ϕ_w	surface work function, V

Subscripts

c,d	convection and diffusion
e	electrons
exp	experimental
h	heavy particles
j	current
L	for a length L_x
w	wall condition
∞	condition at free-stream edge of the boundary layer

by Lees (Ref. 2). Substitution of Camac and Kemp's (Ref. 3) gradients at the wall in the transformed system then yields the following relation of the Nusselt number for a flat plate:

$$Nu_x = -\frac{q_{c,d}\,x}{k_{hw}\,(T_\infty - T_w)} = 0.332\left(\frac{\rho_w\eta_w}{\rho_\infty\eta_\infty}\right)^{0.1}\left(\frac{\rho_w}{\rho_\infty}\,Re_x^{1/2}\right)(\alpha + N_{1w}Sc_w^{2/3}) \qquad (4)$$

where

$$\left.\begin{aligned} Re_x &= \rho_\infty\, U_\infty\, x/\eta_\infty \\ N_1 &= (m_h I_{mi} g_{e_\infty}\eta)/(k_h T_\infty m_e) \\ S_c &= \rho\, D_{amb}/\eta \\ \alpha &= \alpha\,(Pr) \end{aligned}\right\} \qquad (5)$$

The coefficient α shows the effect of Prandtl number on the temperature gradient at the wall. For a constant Prandtl number, it was found by Back (Ref. 4) that $\alpha \propto Pr^{1/3}$, but by exact numerical calculation of the boundary layer of an argon plasma (Ref. 5) at $p = 1$ atm, it was found that $\alpha = 1$ for $T_\infty < 18{,}000$°R and $\alpha = 1.5 - (2.78 \times 10^{-5}\, T_\infty)$ for $T_\infty = 18{,}000$ to 36,000°R. Note that here α is not the ionization fraction. In the present calculation $\alpha = 1$, since $T_\infty \simeq 13{,}000$°R; even though the pressure was less than 1 atm, the ionization fraction was low, and hence the effect of pressure was negligible.

For a wall surface of dimensions L_x and L_y, the rate of total convective and diffusive heat transfer at the wall, as determined from the average heat flux and Eq. (4), is

$$Q_{c,d} = \bar{q}_{c,d}\, L_x\, L_y = -\,2\,Nu_L\, k_{hw}\,(T_\infty - T_w)L_y \qquad (6)$$

where the bar indicates average value and Nu_L is the Nusselt number computed from Eq. (4) at $x = L_x$. The total rate of heat transfer for the entire surface resulting from the current can be found by expressing Eq. (3) in terms of the total current,

$$Q_J = I\left(2.5\,\frac{k_B}{e}\,T_{e_\infty} + \phi_s + \phi_w\right) \qquad (7)$$

According to Cobine (Ref. 6) the work function for a copper surface is $\phi_w = 4.3 \pm 0.2$ V, and according to measurements by Bose and Pfender (Ref. 7) the anode fall potential should be between 0.6 and 2.3 V. Since there is

some uncertainty about these values, a different method is used to evaluate the sum of ϕ_s and ϕ_w as described subsequently.

3. Combined Theoretical and Experimental Results

Comparisons on an absolute basis between theoretical predictions and experimental heat transfer results can only be made if the gas temperature has been evaluated experimentally, which is proposed for the future. Since the gas temperature was not measured in these experiments, the theoretical analysis is used together with the experimental results of one test to determine an approximate value of the gas temperature in the luminous region of the flow (Fig. 22). In addition, the sum of the anode fall potential and the work function $(\phi_s + \phi_w)$ was determined, the approximate origin of the thermal boundary layer was established, and that portion of the heat transferred to the anode which was contributed by the impingement of electrons was deduced. Also discussed are the experimental heat loads to each of the four segments that were in line with the plasma flow for four tests in which the gas velocities were different.

The discussion is confined to trends and comparisons of total heat loads to the individual segments with each segment denoted by an indicator ℓ, or segment number. Introduction of the segment number simplified the theoretical calculations as compared to a determination of the heat flux distribution. It also eliminated the uncertainty in establishing the locations at which the average experimental values of heat flux or heat transfer coefficient of each segment occurred.

Predicted trends in heat load were determined by referencing to the upstream segment (No. 1), which was the anode. It was assumed that for the reference condition the thermal boundary layer originated at the leading edge of this segment ($x_0 = 0$ in Fig. 23). It was also assumed that for both the predicted and reference conditions the free-stream temperature T_∞ was constant along the flow direction. The predicted contribution to the total heat load convection and diffusion, $(Q_{c,d})_{x_0=0}$, for the reference condition and of electron impingement, Q_J, is shown in Fig. 23 as a function of the free-stream temperature T_∞. Four curves for Q_J are shown, each for a different value of the combined anode fall potential and work function $(\phi_s + \phi_w)$ because this is one of the quantities to be determined.

It is evident from Fig. 22, which shows a typical attachment of the discharge, that the thermal boundary layer

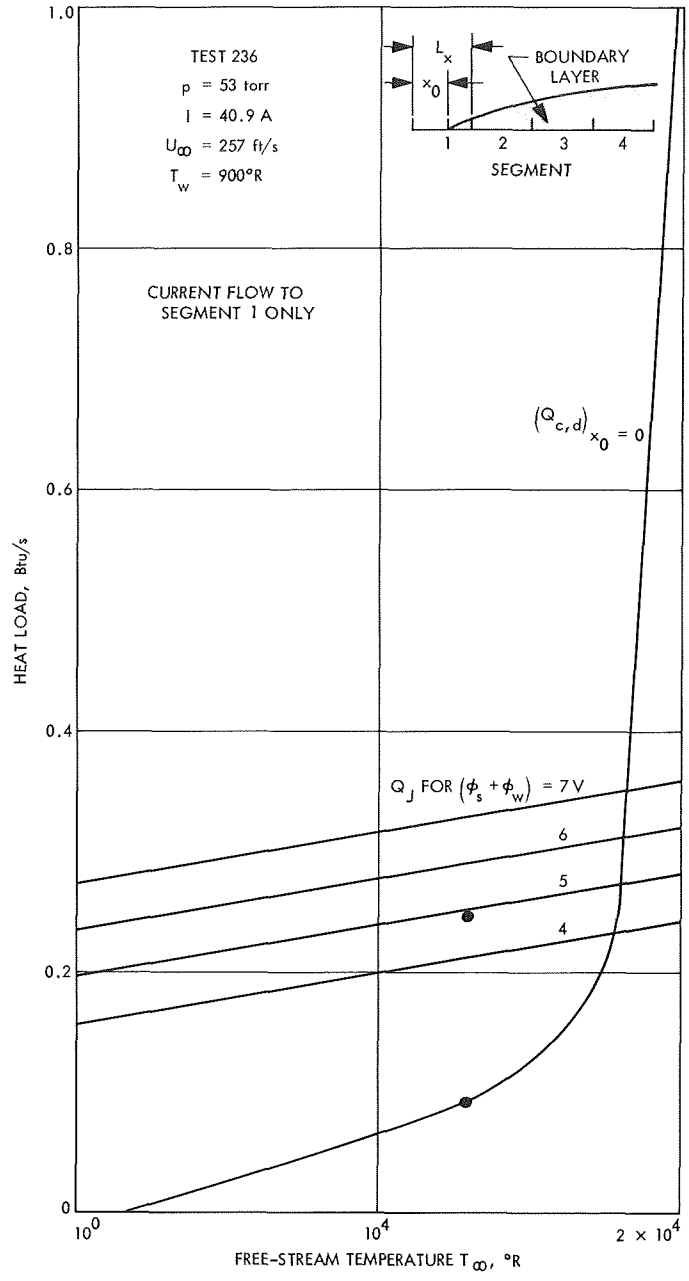


Fig. 23. Theoretical heat loads to the anode (segment 1)

did not originate at the leading edge of segment 1, but appeared to begin closer to the trailing edge instead. Therefore, predictions of the combined convective and diffusive heat loads were made for the three segments downstream of the anode that were in line with the plasma flow for various origins of the thermal boundary layer on segment 1. These heat loads were then normalized to the same reference condition of segment 1 as discussed previously, that is, $(Q_{c,d})_{x_0=0}$. The normalized curves are shown

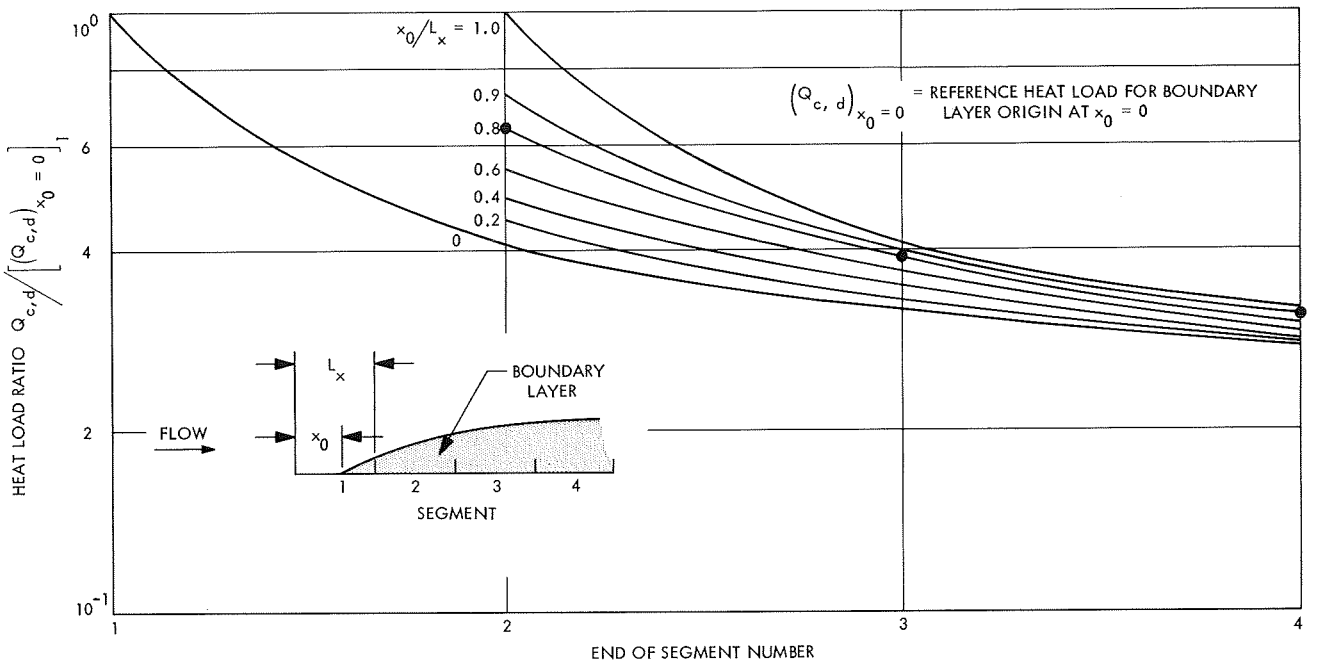


Fig. 24. Theoretical effect of boundary layer origin on convective and diffusive heat loads

in Fig. 24. A simple expression for this normalization resulted from the combination of Eqs. (4) and (6):

$$\frac{[Q_{c,d}]_\ell}{[(Q_{c,d})_{x_0=0}]_1} = \left[\left(\ell - \frac{x_0}{L_x}\right)^{1/2} - \left(\ell - 1 - \frac{x_0}{L_x}\right)^{1/2}\right], \qquad \ell \geq 2 \tag{8}$$

where L_x is the length of each segment, which was 0.25 in.

To determine the location at which the thermal boundary layer originated in one of the experiments (Test 236), the experimental data was plotted in terms of heat load to each segment. The theoretical curve of Fig. 24, which was the best match to the experimental curve, was then selected as a basis for evaluating the approximate origin of the thermal layer. The value of x_0/L_x determined by this method was 0.8; hence, the thermal layer originated on the anode at about 0.20 in. downstream of the leading edge. In addition, for segment 2 the experimental heat load was 0.0512 Btu/s and the corresponding theoretical value of

$$\frac{[Q_{c,d}]_2}{[(Q_{c,d})_{x_0=0}]_1}$$

is 0.64 from Fig. 24. Thus, by assuming that the experimental and theoretical values of $[Q_{c,d}]_2$ are equal,

$$[(Q_{c,d})_{x_0=0}]_1 = \frac{0.0512}{0.64} = 0.080 \text{ Btu/s}$$

Then, by entering this value on the $[Q_{c,d}]_{x_0=0}$ curve of Fig. 23, the free-stream gas temperature T_∞ is found to be about 13,000°R (7220°K).

Based on the equalization of the theoretical and experimental heat loads of segment 2, the combined convective and diffusive heat load to the anode (segment 1) may be expressed as

$$[Q_{c,d}]_1 = \left(1 - \frac{x_0}{L_x}\right)^{1/2} [(Q_{c,d})_{x_0=0}]_1 \tag{9}$$

Equation (9) does not follow directly from Eq. (8) but must be obtained from Eq. (4), since the thermal boundary layer does not completely cover the first segment. The total measured heat load of the anode was 0.283 Btu/s and, by Eq. (9), $[Q_{c,d}]_1 = 0.036$ Btu/s. Therefore, since $Q_J = Q_1 - [Q_{c,d}]_1$, the heat load caused by the current, Q_J, was 0.247 Btu/s or 87% of the total heat load to the anode. Thus, by introducing this value of Q_J into Fig. 23 at the previously determined temperature of 13,000°R, the combined fall potential and work function ($\phi_s + \phi_w$) is found to be 4.9 V, which seems quite reasonable compared with the expected ranges of ϕ_s and ϕ_w stated previously.

The experimental values of the combined convective and diffusive heat loads of the three segments downstream of the anode for Test 236, as well as for three additional tests, are shown in Fig. 25. In the lower part of the figure the data has been collapsed by dividing by $Re_L^{1/2}$, where L refers to the length of one segment (0.25 in.). The close correlation of the data, particularly for segment 2 using the fixed value of L, is an indication that the thermal layer originated at the same location for each of the tests.

From the experimental data it is also possible to determine if the free-stream temperature varied significantly in the flow direction and whether it varied among the tests. By converting Eq. (4) into an expression for the total convective and diffusive heat load of a segment, dividing by $Re_L^{1/2}$ and by those terms that contain the segment indicator ℓ, the following equation may be obtained:

$$\frac{[Q_{c,d}]_\ell}{Re_L^{1/2}\left[\left(\ell - \frac{x_0}{L_x}\right)^{1/2} - \left(\ell - 1 - \frac{x_0}{L_x}\right)^{1/2}\right]} = 2\,k_{hw}\,(T_\infty - T_w)\left(\frac{\rho_w}{\rho_\infty}\right)\left(\frac{\rho_w \eta_w}{\rho_\infty \eta_\infty}\right)^{0.1} \times \left[0.332\,(\alpha + N_{1_w} Sc_w^{2/3})\right] L_y \tag{10}$$

The right side of Eq. (10) is a function of temperature and pressure and, at a given pressure and low values of ionization fraction, is proportional to T_∞^2. For each of the tests the pressure was constant along the flow direction and, furthermore, it was nearly the same for all four tests. Also, the ionization fraction was only about 0.001

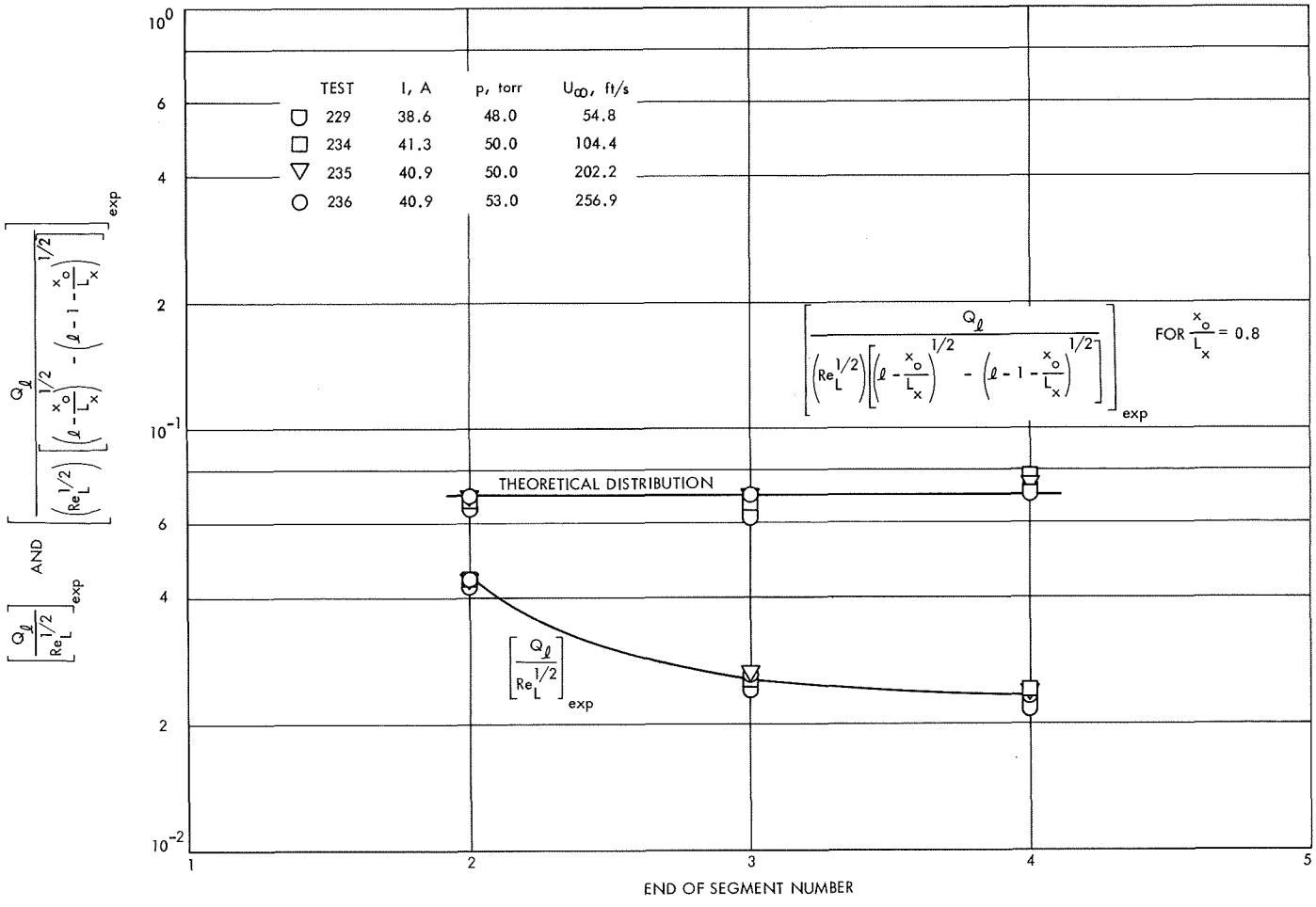


Fig. 25. Experimental results

based on the temperature of 13,000°R. Therefore, if T_∞ were constant along the flow direction, the value of the left side of Eq. (10) should be the same for each of segments 2, 3, and 4. Then, identical values for all tests would indicate that T_∞ was the same for all tests. Both of these conditions were found to exist as is indicated by the data in the upper part of Fig. 25.

4. Summary and Conclusions

Experimental heat transfer results were analyzed in conjunction with a theoretical prediction method for ionized gas flows over a segmented flat surface with an electric field applied to the upstream segment (anode). The tests were conducted in an argon atmosphere at a nominal pressure of 50 torr and a current of approximately 40 A over a gas velocity range of about 55 to 257 ft/s. For one test the free-stream gas temperature associated with the heat transfer was found to be 13,000°R and the combined anode fall potential and work function was 4.9 V. By matching the heat load distribution of the segments with theoretical predictions, it was established that the origin of the thermal boundary layer for all tests occurred at a distance of about 0.8 of the segment length (0.20 in.) from the leading edge of the upstream segment. It was also deduced that the free-stream temperature was constant along the flow direction and that it was the same for all tests. Expressions were presented which were used to evaluate the theoretical heat loads for each of the segments and for collapsing the data.

The theoretical prediction is based on the fact that the temperature gradient at the wall is the same as found from the boundary layer analysis of Bose. Since the current flowing parallel to the downstream segments may alter the temperature distribution, especially for a strongly ionized plasma and for current densities larger than those used in the present case, the method outlined should be used with caution when such conditions exist.

References

1. Bose, T. K., *Laminar Flow Heat Transfer from a Gaseous Plasma at Elevated Electron Temperature in the Presence of Electromagnetic Fields,* Technical Report 32-1447. Jet Propulsion Laboratory, Pasadena, Calif., Oct. 1969.
2. Lees, L., "Laminar Heat Transfer Over Blunt-Nosed Bodies at Hypersonic Flight Speeds," *Jet Propulsion,* Vol. 26, No. 4, pp. 259–269, Apr. 1956.
3. Camac, M., and Kemp, N. H., "A Multiple Temperature Boundary Layer," Paper 63-460, presented at the AIAA Conference on Physics of Entry into Planetary Atmospheres, Cambridge, Mass., Aug. 26–28, 1963.
4. Back, L. H., "Effects of Surface Cooling and Heating on Structure of Low-Speed, Laminar Boundary-Layer Gas Flows with Constant Free-Stream Velocity," Paper 68-HT-23, presented at the AIChE-ASME Heat Transfer Conference and Exhibit Philadelphia, Pa., Aug. 11–14, 1968. Also available as Technical Report 32-1301, Jet Propulsion Laboratory, Pasadena, Calif.
5. Bose, T., *Der Wärmeübergang von Argon- und Stickstoff-Plasma an eine Kalte Wand,* Ph.D. Dissertation. Technical University, Stuttgart, Germany, 1965. (Also appears as Paper DLR-67-17, Proceedings of the German Society of Aeronautical and Space Research on Electric Conversion, Electric Propulsion, and Plasma Flows, Stuttgart, 1967.)
6. Cobine, J. D., *Gaseous Conductors.* Dover Publications, Inc., New York, 1958.
7. Bose, T. K., and Pfender, E., "Direct and Indirect Measurements of the Anode Fall in a Coaxial Arc Configuration," *AIAA J.,* Vol. 7, No. 8, pp. 1643–1645, Aug. 1969.

XX. Liquid Propulsion

PROPULSION DIVISION

A. Advanced Combustion Device Development,

R. W. Riebling

1. Introduction

A number of research programs have been conducted by the Liquid Propulsion Section to evolve injector and thrust chamber design criteria applicable to the development of rocket engines for unmanned spacecraft, using the space-storable propellant combination of oxygen difluoride (OF_2) and diborane (B_2H_6). One objective of the present program, "Advanced Combustion Device Development," is the application of these criteria to the development of a prototype OF_2/B_2H_6 flight propulsion rocket engine. One likely application for such an engine is the JPL space-storable propellants propulsion module, currently being actively studied for planetary orbiting missions around Mars, Venus, or Jupiter. The adopted technical approach emphasizes the application of existing technology, rather than the generation of new technology, with the intent of exposing any problems encountered in an actual application. These will be solved as they arise, or made the subject of additional new technology programs, as appropriate. Another program objective is the demonstration of the basic soundness of currently accepted design concepts, in terms of performance and stability levels, long-duration durability, and inherent reliability.

Of the problem areas uncovered by the technology effort thus far, perhaps the most stringent is that of guaranteeing the structural integrity of an injector/chamber combination in very-long-duration firings under the severe environmental conditions imposed by the OF_2/B_2H_6 combustion process. This involves suppressing throat erosion, chamber streaking, and injector melting at flame temperatures on the order of 7000°F, in the presence of extremely reactive and corrosive products of combustion. The adiabatic wall temperature in the chamber and at the throat must be minimized without incurring undue performance penalties, and heat flux to the injector and chamber must be reduced. Materials are required that are resistant to mechanical erosion by liquid or possibly solid products in the combustion gas stream, and to chemical attack as well. Even at wall temperatures as low as 3500°F, many materials are chemically attacked by the combustion products, which contain such corrosive species as fluorine, oxygen, and boron.

The thrust chamber design criteria evolved to date have been based on the concept of a passively cooled chamber, combined with an injector having a controlled gradient in mixture ratio. That is, the injector includes a central core region, containing a large fraction of the

total mass flow and operating at or slightly above the mixture ratio for maximum performance, and an outer or "barrier" zone, containing a smaller proportion of the total mass flow and operating at a much-reduced mixture ratio. The fuel-rich conditions prevailing in this barrier zone should reduce both the adiabatic wall temperature and the concentrations of chemically undesirable species in the region immediately adjacent to the walls, thus helping to maintain the chamber's structural integrity. Nickel has emerged as the most durable injector material with fluorinated oxidizers such as OF_2, and carbonaceous or graphitic materials have been reported to be the most promising candidates for use as chamber liners and throat inserts in chambers thermally insulated by ablative or other materials.

Accordingly, the technical approach adopted to develop a flight-prototype OF_2/B_2H_6 rocket engine has been based on a passively cooled chamber, with a carbonaceous or graphitic inner liner, and an injector made of nickel which incorporates some degree of barrier zone protection. In the event that the rocket engine must be "buried" within a spacecraft, its outside wall temperature and heat flux per unit area may not be permitted to exceed certain values. Therefore, an additional requirement imposed upon the engines currently being developed is that they be adequately thermally insulated.

Four competing versions of this general design concept are being evaluated. One was designed and fabricated at JPL, while the other three were made for JPL by industrial contractors. For reasons of economy, and also because there may ultimately be a need for an engine of this size, the initial evaluation is being conducted at the 200-lbf thrust level. To assure commonality, all chambers were manufactured in compliance with a single JPL control drawing, and will be evaluated with interchangeable, JPL-furnished injectors. Based on performance and durability in 1000-s firing tests, the most promising chamber concept will be selected for final optimization and the development of prototype flight hardware at the 1000-lbf vacuum thrust level.

One thrust chamber assembly was described in SPS 37-57, Vol. III, pp. 187–192. This SPS article will describe the design and fabrication of the injectors and two other chamber concepts.

2. Injectors

a. General requirements. In accordance with the emphasis placed by the technical approach on maximum utilization of existing technology, the 200-lbf flight prototype injectors being designed and fabricated are generically similar to designs which have evolved under a JPL-managed NASA advanced technology contract (Ref. 1). For example, they feature flow stratification into a central core and a surrounding barrier zone (as mentioned in the Introduction), have multiple self-impinging doublet injection elements arranged in a radial pattern (to minimize direct wall impingement of spray fans), and are constructed chiefly of high-purity nickel (the injector material found to be most compatible with the OF_2/B_2H_6 combustion environment).

In addition to the above-mentioned state-of-the-art design criteria, however, several other requirements were imposed by the flight prototype status of the injectors. One of these was a hump-type mass flux distribution (Ref. 2). Such a distribution concentrates injected mass flux in an annular band bracketing the locus of half-radius points on the injector face, and is frequently found to reduce the likelihood of the first tangential and first radial modes of high-frequency combustion instability. In addition, it generally causes recirculation patterns, resulting in the flow of combustion gases radially outward across the injector face (Ref. 3). While sometimes harmful to the structural integrity of wall materials in chambers not protected by a fuel-rich barrier zone, these radial "winds" may favorably alter the undesirable solids deposition patterns characteristic of previous, non-hump-distributed, OF_2/B_2H_6 injectors (Ref. 4). Other requirements included interchangeability with all four thrust chamber designs, a compact envelope, and the elimination of all internal seals. Finally, the injectors were to incorporate provisions for controlled, reproducible hydraulics, to the extent possible in a design of flight volume and weight.

b. Specific designs. Two similar 200-lbf injectors were designed and fabricated. Their key dimensions and flow proportioning are summarized in Table 1. The injector face patterns are shown in Fig. 1, and their manifolding arrangement is illustrated by Fig. 2. Both injectors were designed for a total propellant flow rate of 0.55 lbm/s and a maximum core mixture ratio of 3.85, assuming $-100°F$ B_2H_6 and $-315°F$ OF_2. The stagnation chamber pressure will depend on the actual degree of core/barrier interaction, but the nominal value is 100 psia. The designs will be described with the aid of Table 1 and Figs. 1 and 2.

The Mod I injector, with 68% of the total fuel flow in its outer ring of fuel doublets, and a correspondingly

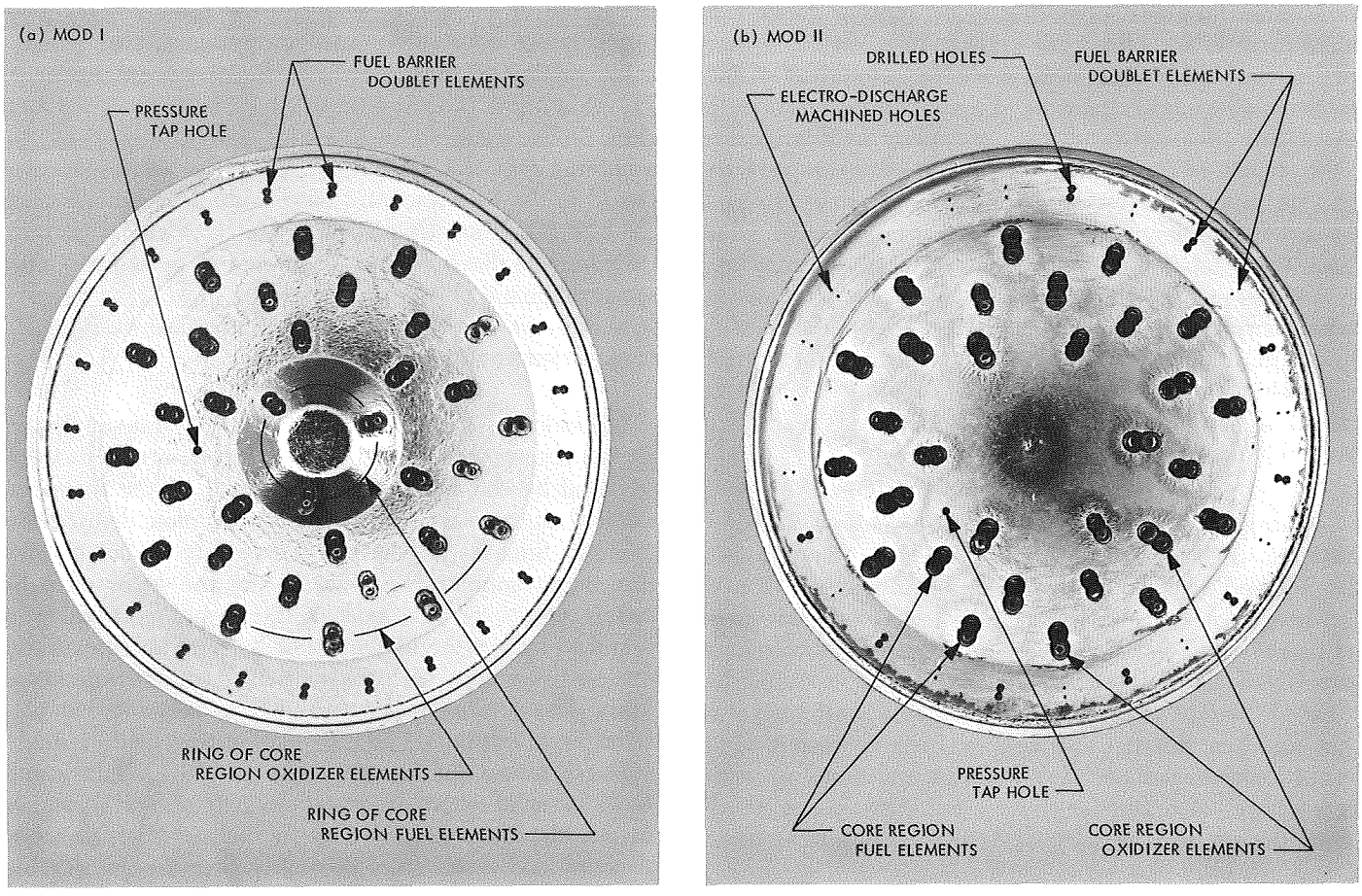


Fig. 1. Element arrangements of 200-lbf impinging-jet injectors

lower (1.23) overall mixture ratio, is designed to provide a greater degree of barrier protection than the Mod II version. The latter has only 41% of the total fuel flow in its outer ring, and a higher (2.27) overall mixture ratio. The maximum possible core mixture ratio for both injectors (assuming no mixing between the core and the fuel from the outer ring) is 3.85. This value should maximize core characteristic velocity c^* at a chamber pressure of 100 psia. The "barrier zone" will be formed by the interaction of the oxidizer-rich core sprays and the fuel sprays from the outer ring. Since it is not possible to predict the degree of this interaction under combustion conditions, it would be fictitious to assign an average mixture ratio to the barrier region; in all probability, however, a mixture ratio gradient will exist across the barrier zone, from a value of approximately 3.85 at its boundary with the core, to one approaching zero at the chamber walls.

The Mod I injector consists primarily of alternating rings of fuel and oxidizer doublets, while the Mod II injector contains alternating fuel and oxidizer doublets within each ring. Thus, the Mod I design maximizes the spacing between the long axes of the elliptical cross sections of fuel and oxidizer spray fans, but the Mod II concept minimizes (reduces to zero) this spacing. It has been shown (Ref. 5) that fan spacing is a key variable affecting the degree of mixing in self-impinging doublet injectors, so the Mod II injector would be expected to provide better mixing and higher combustion efficiency in its core region than the Mod I. This is borne out by the values of the mixing factor E_m reported in Table 1 (see also *Subsection 2d*, Cold-Flow Spray Analysis). These attributes of barrier protection and core performance will combine in such a way as to make the chamber and nozzle heat fluxes expected from the Mod II injector considerably higher than those from the Mod I. The Mod II should provide higher combustion efficiencies and a more severe chamber environment than the Mod I, thus providing comparative information which will hopefully lead to an optimized injector pattern for the subsequent 1000-lbf thrust chamber assemblies.

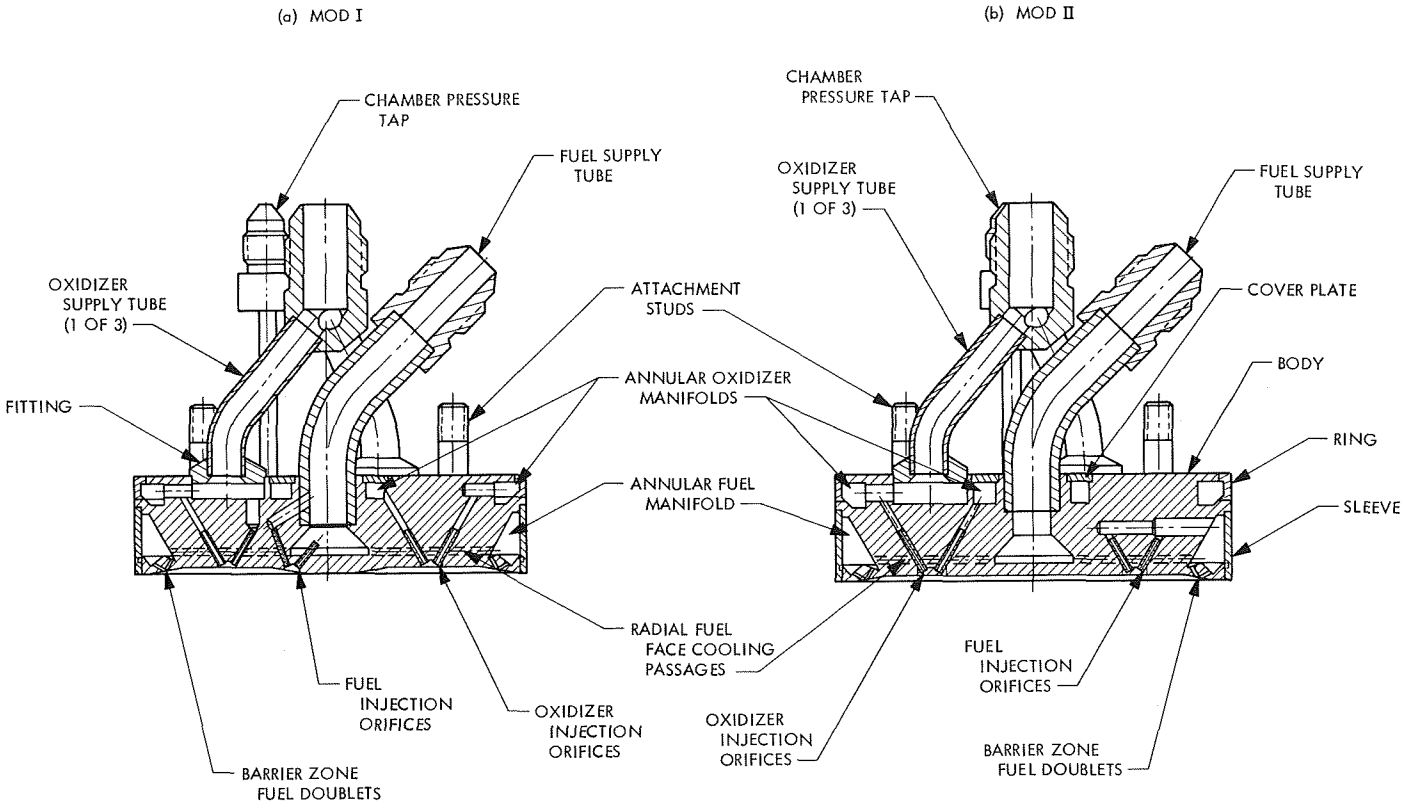


Fig. 2. Manifolding of 200-lbf impinging-jet injectors

Table 1. Summary of injector parameters

Parameter	Mod I injector	Mod II injector
Oxidizer doublet elements, number	24	15
Oxidizer orifice diameter, in.	0.0170	0.0215
Core fuel doublet elements, number	9	15
Core fuel orifice diameter, in.	0.0151	0.0166
Outer ring fuel doublet elements, number	24	24
Outer ring fuel orifice diameter, in.	0.0135	0.0110
Oxidizer flow area, in.^{2}	1.09×10^{-2}	1.09×10^{-2}
Core fuel flow area, in.^{2}	0.326×10^{-2}	0.66×10^{-2}
Outer ring fuel flow area, in.^{2}	0.695×10^{-2}	0.461×10^{-2}
Core fuel flow rate, lbm/s	0.0785	0.0989
Outer ring fuel flow rate, lbm/s	0.1675	0.0691
Oxidizer flow rate, lbm/s	0.304	0.382
Total propellant flow rate, lbm/s	0.550	0.550
Maximum core mixture ratio (O/F)	3.85	3.85
Overall mixture ratio (O/F)	1.23	2.27
Fuel flow in outer ring, wt %	68	41
Mixing factor E_m, 3.0 in. from face	0.7257	0.7911
Dry weight, lbm	0.99	1.02

The manifolding schemes of both injectors (Fig. 2) are virtually identical; the apparent differences in Fig. 2 result from the positions at which the sections were taken. Fuel enters each injector through a large-diameter supply tube, and then flows radially outward at about 29 ft/s through a series of cooling passages close to the face. It then enters an annular fuel manifold which feeds the fuel doublets of the outer ring directly, and the fuel doublets in the core region via the step-drilled feeder passages. Oxidizer is distributed to two concentric annular manifolds on the backside of the injector via the three-branched supply tubes. These manifolds, in turn, feed all the oxidizer doublets in the core.

All orifices in the core region have length-to-diameter ratios L/d of 15:1 and contoured entrances (radius of curvature equal to orifice inside radius). In addition, the first five diameters of length beyond the entrance of each core orifice are threaded with National Miniature Standard screw threads. These measures were taken to provide the equivalent of fully developed turbulent flow from each orifice, to assure predictable and reproducible flow characteristics from one element to the next. It has been shown (Ref. 6) that such an orifice design can effectively simulate the hydraulic characteristics of the

much longer ($L/d \cong 100$) orifices generally required for fully turbulent velocity profiles. Spatial constraints required that simple drilled holes, 5 diameters in length, be employed for barrier fuel injection.

c. Fabrication sequence. The individual components were fabricated by conventional machining techniques, inspected, and joined by vacuum furnace brazing. Prior to brazing, all parts were vapor-degreased in hot trichlorethylene and ultrasonically cleaned in a fluorocarbon solvent.

The brazing was accomplished in two steps. In the first step, the core orifices were inserted into corresponding holes drilled in the injector body, staked in place, and brazed at 1800°F with an alloy containing 82 wt % gold and 18 wt % nickel, applied externally in slurry form. A typical oxidizer doublet element after brazing is shown in the photomacrograph of Fig. 3, where the braze fillet is also visible. These joints were pressure-checked to 250 psig in a special fixture, and found to be leak-tight.

In the second step, the ring, sleeve, and cover plate (Fig. 2) were joined to the body, thereby completing the manifolds, and the supply tubes, studs, and fittings were attached. The same alloy and temperature were used. A completed injector is shown in Fig. 4. No leaks were detected upon subsequent pressure-checking.

Several orifices in the fuel barrier ring of the Mod II injector were found to be oversize, resulting in fuel spray fans that were canted away from the walls at unacceptably steep angles. These doublets were welded shut, and new elements electric-discharge machined (EDM) immediately adjacent to them. Twelve additional doublets were drilled between the original twelve during the same operation (to bring the total to the 24 shown in Table 1). Those drilled by the EDM operation required no spot facing, which accounts for their different appearance in Fig. 1a.

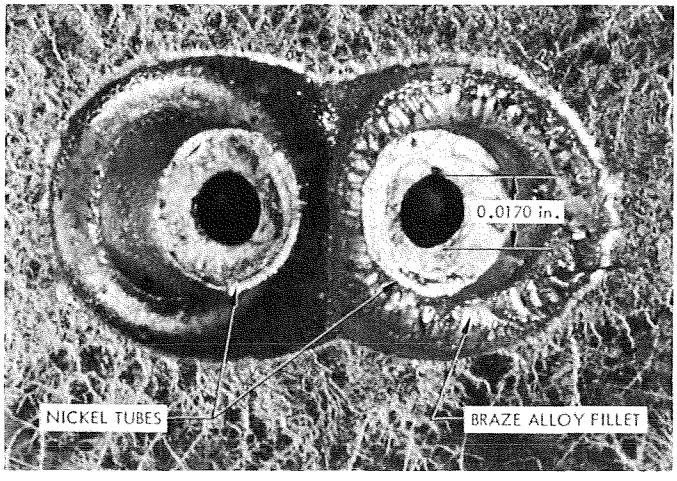


Fig. 3. Photomacrograph of typical oxidizer doublet element

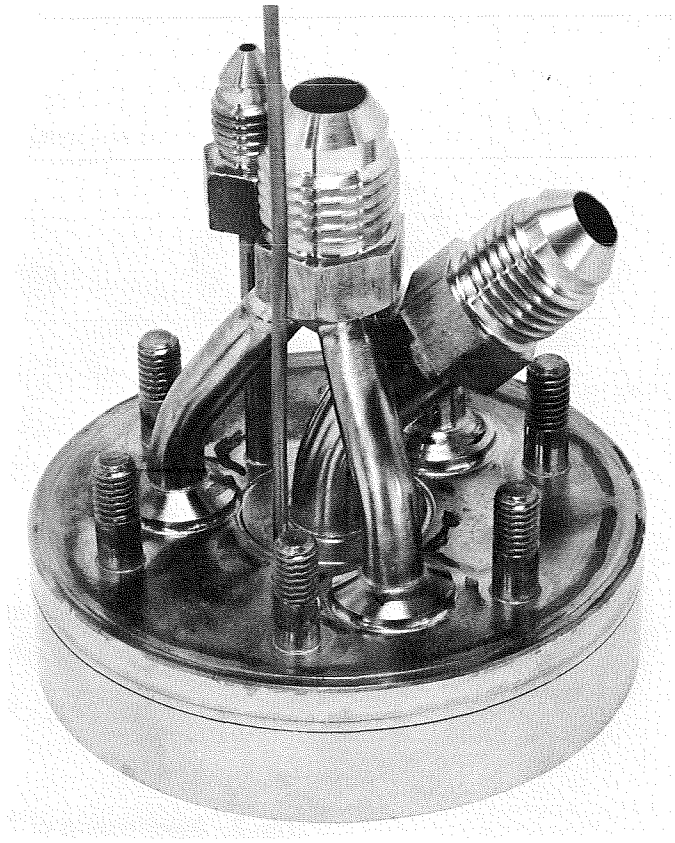


Fig. 4. Completed injector assembly

d. Cold flow spray analysis. The injectors were flowed with nonreactive, immiscible propellant simulants to determine their mass and mixture ratio distributions under nonreactive conditions. OF_2 was simulated by water, and B_2H_6 was simulated by *n*-hexane. The simulants were passed through the injector at volumetric flow rates equivalent to the volumetric flow expected with the actual propellants, and collected in standard spray collection devices at the JPL Hydraulics Laboratory. The sprays were sampled in a plane 3.0 in. downstream of the injector faces. Simulant densities were corrected to propellant densities during data reduction, so the resulting distributions represent propellant distributions in the absence of disruptive combustion effects.

The results are summarized in Figs. 5 and 6. Figure 5a shows the mass flux distribution for the Mod I injector, and indicates that the desired mass flux concentration at the half-radius location was achieved. The corresponding distribution of local mixture ratios is shown in Fig. 5b. There is a gradual gradient from a mixture ratio of 0.1 at the outer periphery of the spray to values in excess of 3 in the main core. Superposition of Figs. 5a and 5b reveals that the low mixture ratios are concentrated in the low mass flux regions at the outer extremities of the spray. The greater degree of barrier protection afforded by this injector (compared to the Mod II version) is obtained at the expense of less efficient spray mixing, as indicated by the lower value of the mixing factor E_m reported in Table 1.

Mass and mixture ratio distributions for the Mod II injector are shown in Fig. 6. The mass flux concentration near the half-radius position is again apparent, as seen in Fig. 6a. There is a steeper mixture ratio gradient across the barrier zone from values less than 0.1 (essentially pure fuel) at its outer periphery to values ranging between 3 and 4 in the main core, as planned (Fig. 6b). Only rarely does the mixture ratio exceed 5 in the core, as indicated by the "G" contours. With less barrier protection, the mixing factor E_m is higher (Table 1). Although not perfectly symmetrical, these patterns do indicate that the desired mass and mixture ratio distributions have been closely approximated and that an appreciable degree of fuel-rich barrier protection has been afforded.

3. Thrust Chamber Assemblies

a. General requirements. The basic ground rules and nominal operating conditions for which the chambers were designed and fabricated are shown in Table 2. The relatively large contraction area ratio is a result of the large injector diameter (2.6 in.) required to accommodate the number of injection elements necessary to assure an adequate mixture ratio gradient. Lower contraction ratios are conducive to improved secondary atomization and hence somewhat higher combustion efficiencies with this propellant combination, and would be preferred for larger scale engines. The expected slight effect on vaporization efficiency was offset in the present case by the use of a somewhat longer than usual chamber characteristic length L^* of 20 in.

All the chambers have an expansion area ratio of 5.0 and are provided with an external lip to interface with a straight-tube diffuser at the JPL test facility. This large expansion ratio (for ground-level tests) was necessary to provide a large enough throat and divergent nozzle heat transfer area to assure the meaningful evaluation of each chamber concept. In the absence of an altitude chamber, the diffuser will assure that these nozzles flow full during the firings.

Table 2. General Requirements of 200-lbf OF_2/B_2H_6 rocket engines

Parameter	Requirement
Core mixture ratio	3.85
Nominal chamber pressure, psia	100
Total propellant flow rate, lbm/s	0.55
Throat diameter, in.	1.162
Chamber inside diameter, in.	2.6
Chamber characteristic length, in.	20
Contraction area ratio	5.0
Expansion area ratio	5.0
duration, s	1000 (continuous)
Restart capability	multiple
Throttling capability	none required
Maximum allowable outside wall temperature, °F	600

The maximum allowable outside wall temperature shown in Table 2 is based on the potential requirement for a thrust chamber completely enclosed within a spacecraft.

b. Marquardt[1] thrust chamber concept. A complete thrust chamber assembly is depicted in Fig. 7, and its physical characteristics are summarized in Table 3. The design concept will be discussed with the help of Figs. 7–9.

The design features a free-standing, monolithic chamber liner made of Carb-I-Tex 713[2] (Fig. 8). This material is made from three-circuit, filament-wound graphite yarn, which has been resin-impregnated and regraphitized. The inner surfaces of the liners have been infiltrated to a nominal depth of 0.050 in. with in-situ-deposited pyrolytic graphite. The primary seal against chamber pressure is the tapered, lapped joint between the pyrolytic graphite and the stainless steel chamber adaptor flange; the O-ring seal shown in Fig. 7 is redundant. The liner is held in place against the adaptor flange by the chamber retainer flange.

[1]Manufactured by The Marquardt Corp., Van Nuys, Calif.

[2]Manufactured by The Carborundum Co., Niagara Falls, N.Y.

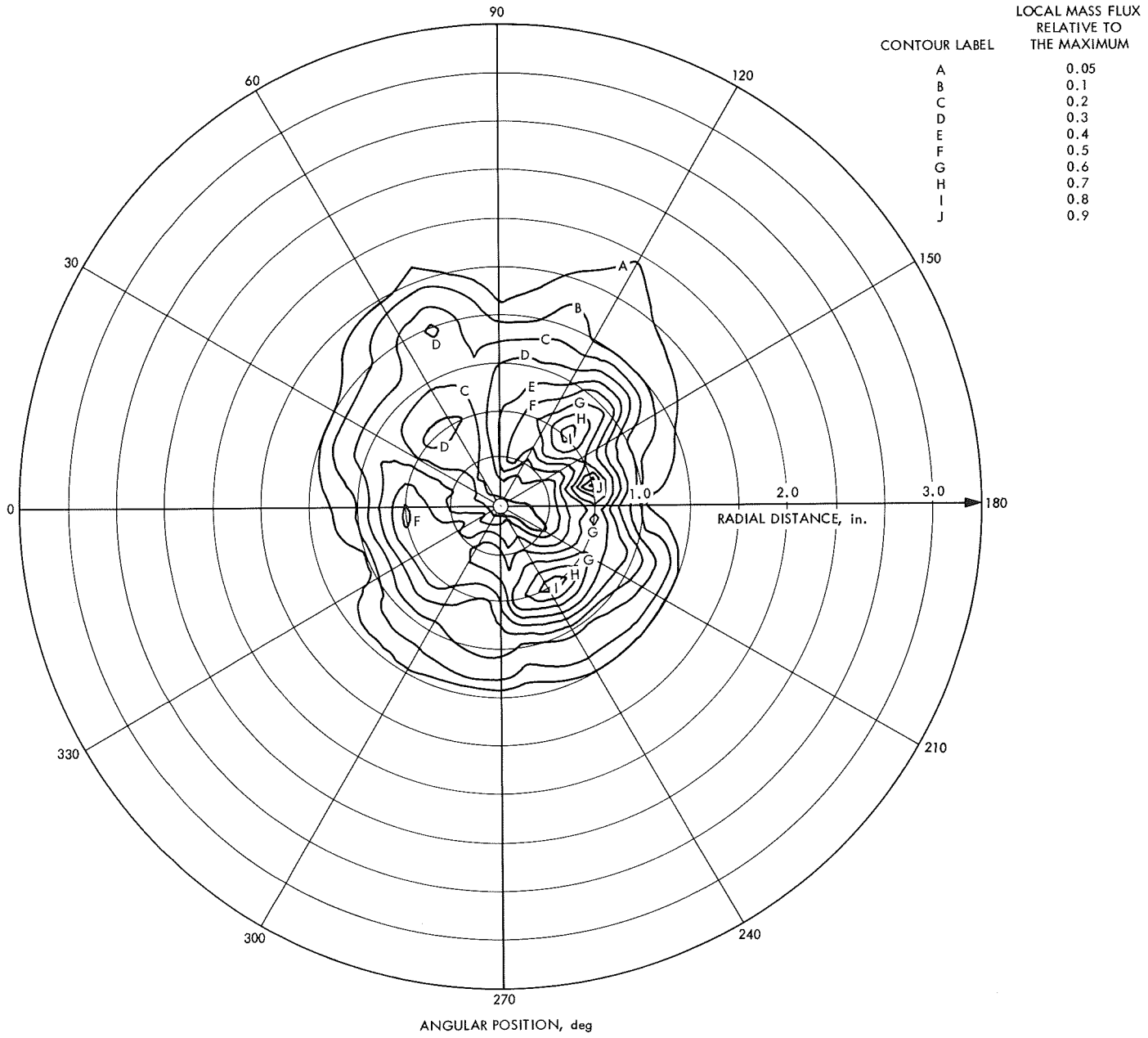


Fig. 5. Contour plots for Mod I injector: (a) normalized local mass concentration, (b) local mixture ratio

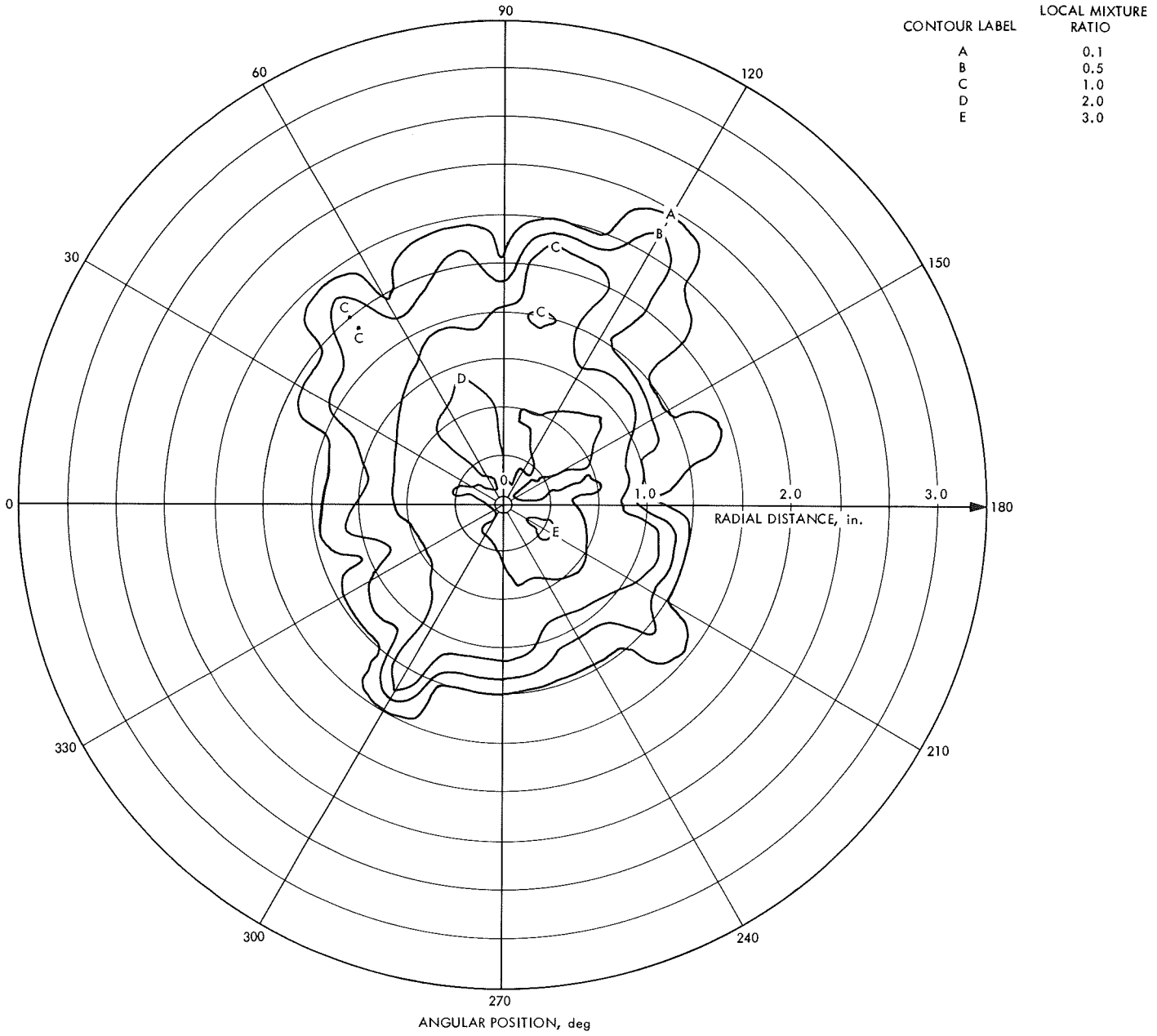

CONTOUR LABEL
LOCAL MIXTURE RATIO
A 0.1
B 0.5
C 1.0
D 2.0
E 3.0
90
60
120
30
150
0
180
330
210
300
240
270
0
1.0
2.0
3.0
RADIAL DISTANCE, in.
ANGULAR POSITION, deg

Fig. 5 (contd)

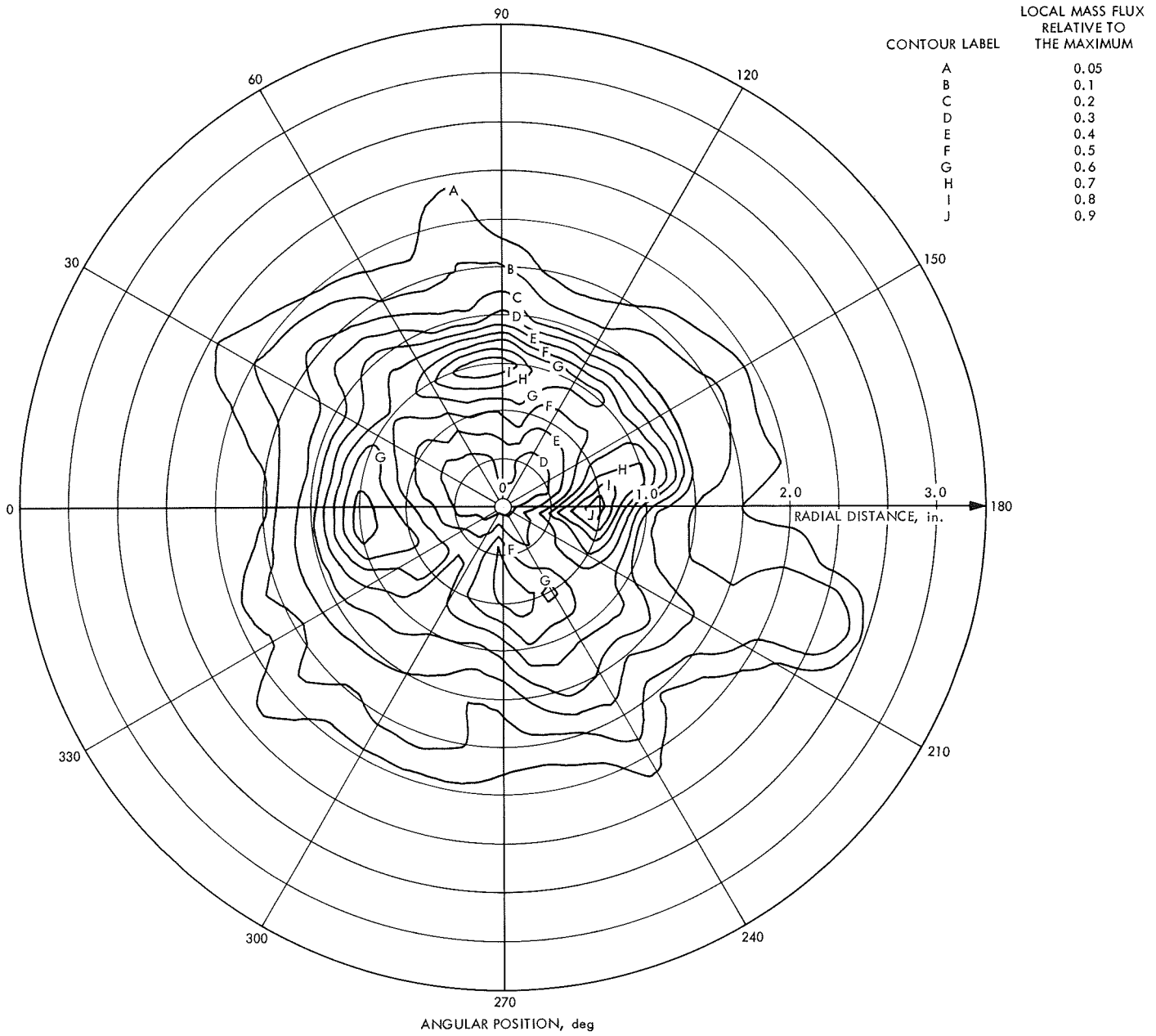


Fig. 6. Contour plots for Mod II injector: (a) normalized local mass concentration, (b) local mixture ratio

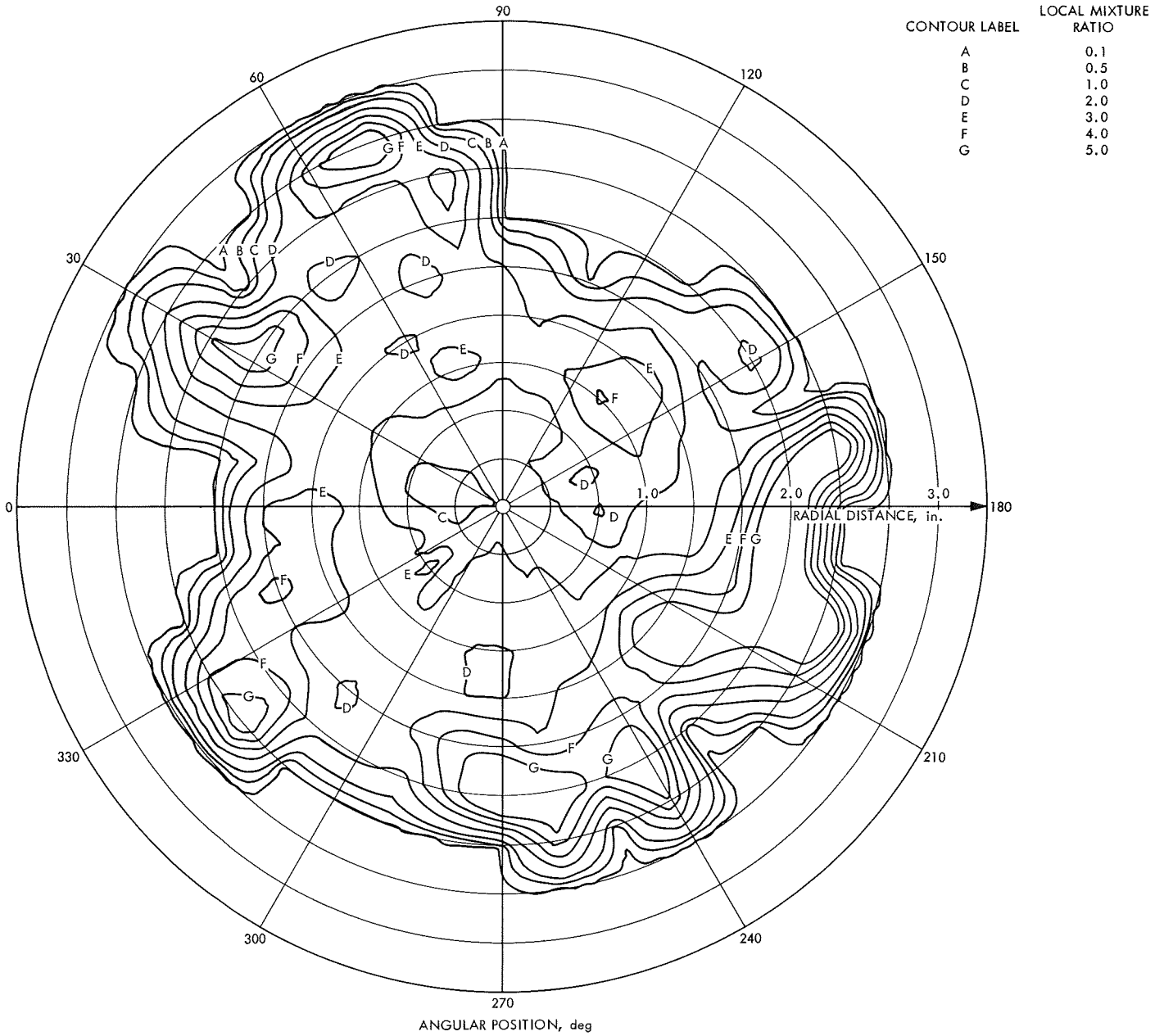

CONTOUR LABEL
LOCAL MIXTURE RATIO
A 0.1
B 0.5
C 1.0
D 2.0
E 3.0
F 4.0
G 5.0
90
60
120
30
150
0
180
330
210
300
240
270
1.0
2.0
3.0
RADIAL DISTANCE, in.
ANGULAR POSITION, deg

Fig. 6 (contd)

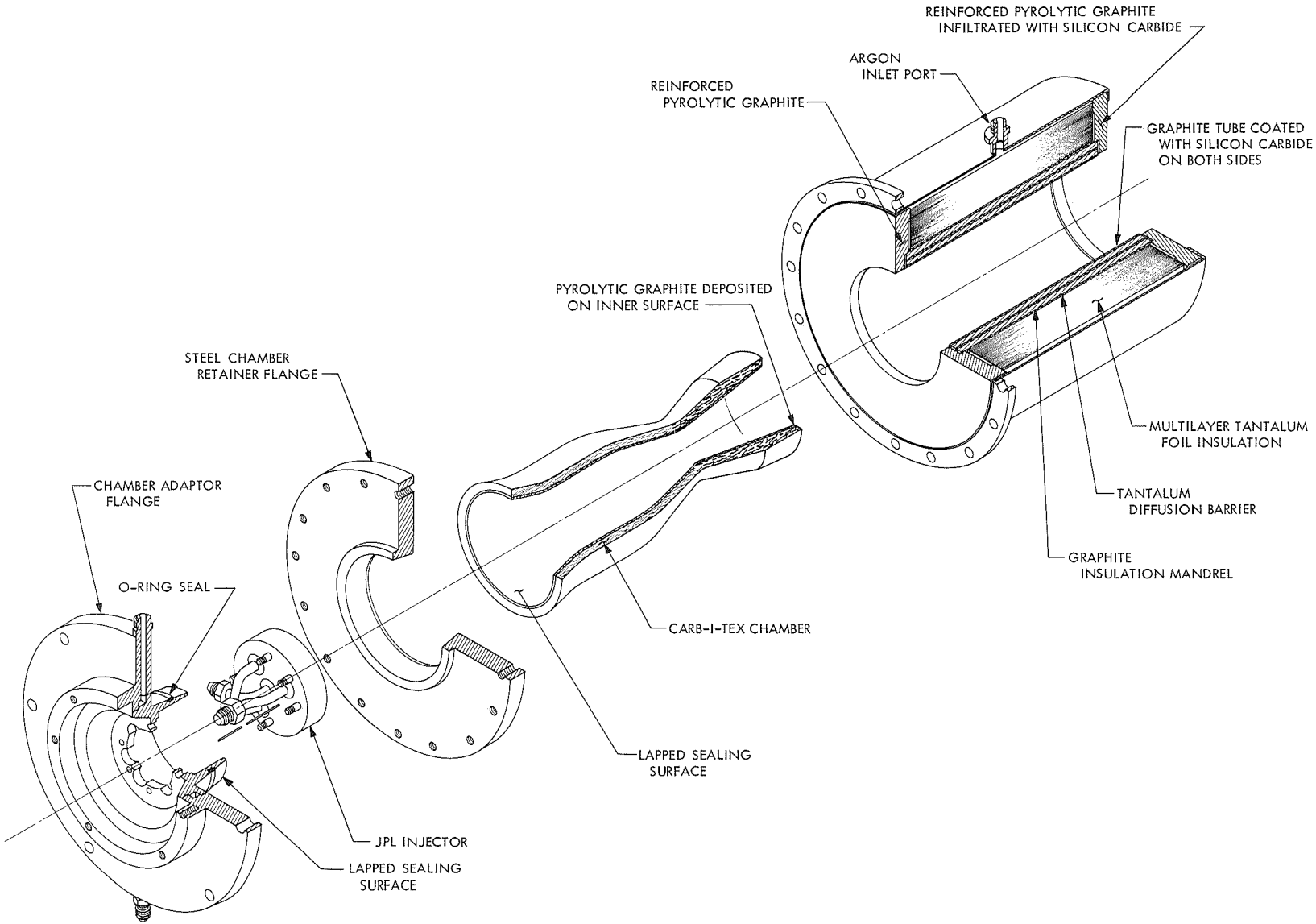


Fig. 7. Exploded view of 200-lbf OF_2/B_2H_6 thrust chamber (Marquardt Corp.)

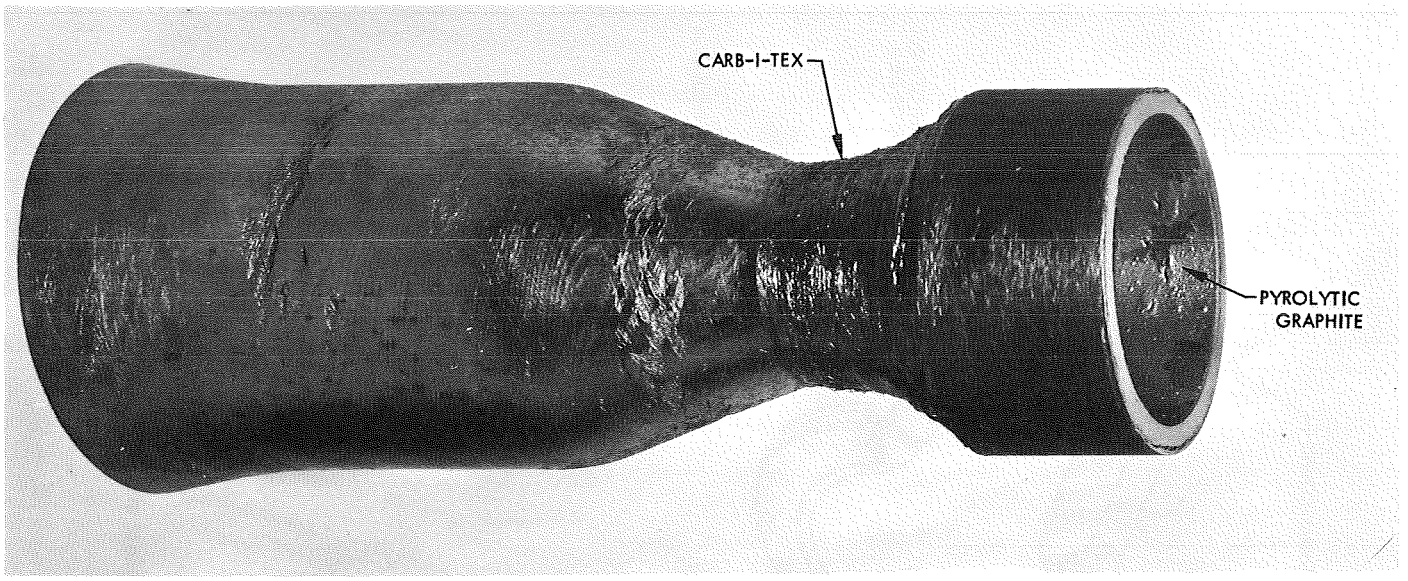


Fig. 8. Chamber liner (Marquardt Corp.)

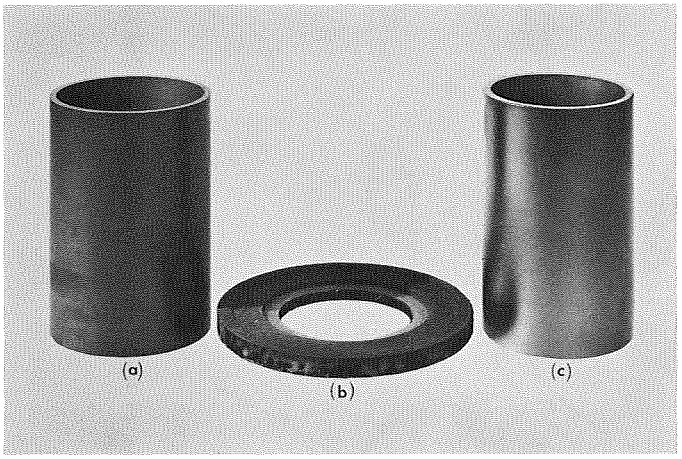


Fig. 9. Marquardt Corp. insulation package: (a) graphite mandrel, (b) SiC end plate, (c) SiC-coated tube

The insulation package surrounds the chamber liner and is completely removable. That is, with adequate protection from oxidation by the air, the chamber liner could be fired in a radiation-cooled mode. The foundation of the insulation package is a Carbone 2239[3] graphite tube (Fig. 9c), coated inside and out with 0.050-in. thick, vapor-deposited Marquardt RM-005 silicon carbide to provide oxidation resistance during ground-level testing. The insulation, which consists of multiple layers of tantalum foil, is wrapped around a Carbone 2239 graphite mandrel (Fig. 9a) surrounding the previously mentioned tube. A tantalum diffusion barrier (40 layers of ½-mil-thick tantalum foil) has been placed between the graphite mandrel and the tantalum insulation; its purpose is to preferentially react with the graphite, forming tantalum carbide at the elevated temperatures of engine operation, thereby protecting the first few layers of ½-mil-thick tantalum insulation from degrading due to carbide formation. The end closures (Fig. 9b) of the insulation package are made of reinforced pyrolytic graphite (RPG).[4] This is a material made from graphite felt infiltrated with pyrolytic graphite. The aft end closure has, in addition, been infiltrated with silicon carbide to resist oxidation by air aspirated past it into the inlet of the diffuser which will mate with the nozzle exit during ground testing. A stainless steel outer shell encloses the entire package and joins it to the chamber retainer flange.

The ground test configuration shown in Fig. 7 requires a low-pressure argon gas purge during firing to eliminate air from the system and prevent oxidation of the tantalum and graphite components. This, of course, would not be necessary in the hard vacuum of space operation. Marquardt elected to use a purged open system, rather than housing the insulation in an evacuated can, and the ground-test version has a thicker wrapping of tantalum foil than would be required in space, because of the increased conductivity resulting from the presence of the gas. The greater thermal mass of the ground-test engine relative to a flight model provides for greater heat storage

[3]Product of Le Carbone Lorraine, 45 Rue Acacias, Paris, France.

[4]Product of Super Temp Co., Santa Fe Springs, Calif.

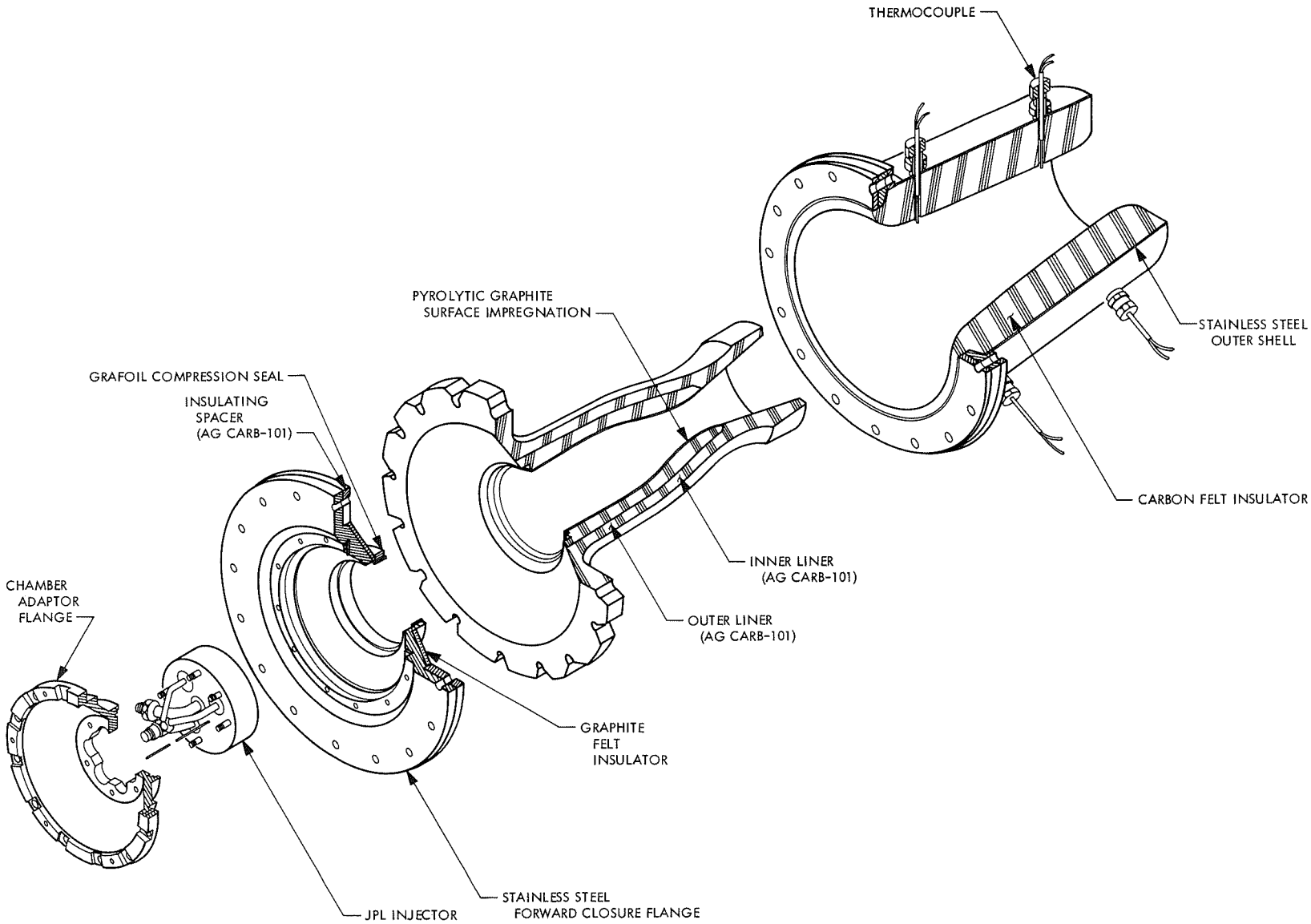


Fig. 10. Exploded view of Aerojet-General Corp. 200-lbf OF_2/B_2H_6 thrust chamber assembly

during a firing and, accordingly, greater post-test soakback. Thus, the chamber adaptor flange in the present version is water-cooled. This would not be the case with a flight configuration.

c. Aerojet-General Corp.[5] thrust chamber concept. The Aerojet-General Corp. furnished two modifications of their basic thrust chamber concept. These will be referred to in this article as Mod I and Mod II. The Mod I assembly (Aerojet-General part 1156905) has a tapered combustion chamber and a black teflon ablative insulator, while the Mod II assembly (Aerojet-General part 1156929) has a cylindrical combustion chamber and a carbon felt insulator. In this way, two different chamber geometries and insulation systems may be compared under otherwise identical conditions. In all other respects, the two chambers were identical. The physical characteristics of the chambers are summarized in Table 3. The design concept will be discussed with the aid of Figs. 10–12.

Table 3. Physical characteristics of thrust chamber assemblies (less injectors)

Parameter	Marquardt	Aerojet-General Mod I	Aerojet-General Mod II
Overall length, in.	10.0	9.88	9.88
Maximum outside diameter, in.	8.0	7.95	7.95
Dry weight, lbm	33	24	12
External volume, in.^{3}	327.9	305	305

The design features a free-standing, two-piece chamber liner made of AG Carb-101[6] (Fig. 11). This material is ordinarily made from graphite yarn (wrapped at an angle of 45 deg to the chamber's longitudinal axis) which has been resin-impregnated and regraphitized. In the present case, additional resin-impregnation and regraphitization per the procedures developed in Ref. 7 was applied to reduce the gas permeability of the liners to an acceptable level. The inner surfaces of the liners have been infiltrated to a nominal depth of 0.10 in. with in-situ-deposited pyrolytic graphite. The primary seal against chamber pressure is a Grafoil[7] compression gasket (Fig. 12) positioned between the insulating spacer and the serrated flange of the outer chamber liner. The stainless steel forward closure flange is thermally insulated from the flange of the outer liner by graphite felt and AG Carb-101 spacers (Fig. 12). The JPL injector is held in place by the chamber adaptor flange (Fig. 10). The chambers could be fired in this configuration (less the insulation system) if provisions were made to protect the exterior surfaces from oxidation by the surrounding air during ground testing.

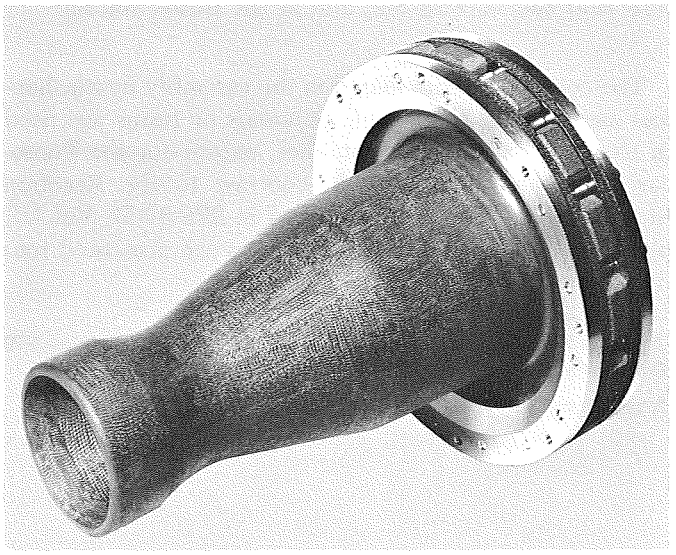


Fig. 11. Assembled Mod II thrust chamber less exterior insulation (Aerojet-General Corp.)

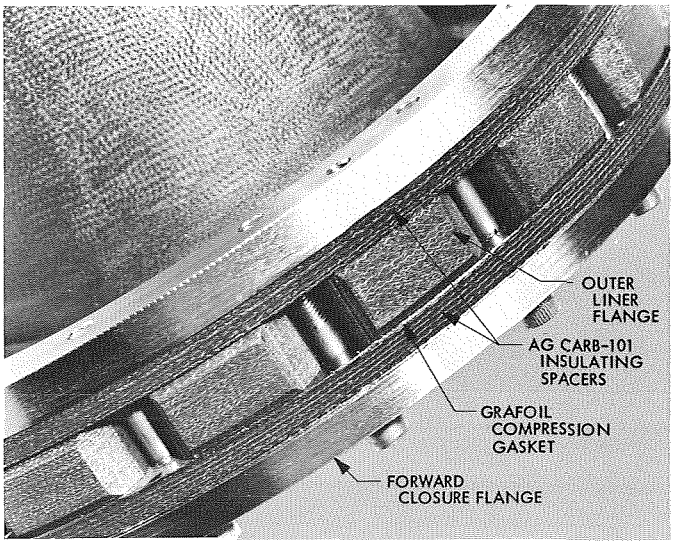


Fig. 12. Details of flange assembly (Aerojet-General Corp.)

[5]Manufactured by Aerojet-General Corp., Sacramento, Calif.

[6]Manufactured for Aerojet-General Corp. by San Rafael Plastics Co., 97 Jordan Street, San Rafael, Calif.

[7]A product of Union Carbide Corp., Carbon Products Division, 270 Park Avenue New York, N.Y.

The insulation package surrounds the chamber liner and is completely removable. In the Mod II assembly (shown in Fig. 10), the insulation consists of multiple layers of ¼-in.-thick carbon felt (grade VDG)[8], formed by rolling up a continuous mat around the chamber liner. In the Mod I assembly, the insulator has been machined from a solid billet of black TFE fluorocarbon[9] ablative material. A stainless steel outer shell encloses the entire package and joins it to the forward closure flange. Both thrust chamber assemblies are provided with thermocouples to monitor the inner-liner backside wall temperature near the injector end and at the throat station.

The components are joined in the assembly by mechanical means only. No organic polymers or resins are used in the thrust chamber assemblies, except for the fluorocarbon insulator. This is in contrast to the Reaction Motors thrust chambers, described in SPS 37-57, Vol. III, pp. 187–192, wherein the components are cemented into the assembly with organic adhesives.

[8]A product of Union Carbide Corp., Carbon Products Division, 270 Park Avenue, New York, N.Y.

[9]A product of Raybestos-Manhattan Co., Fullerton, Calif.

References

1. *Chamber Technology for Space-Storable Propellants, Third Interim Report,* Contract NAS7-304, Report R-7073. Rocketdyne, Canoga Park, Calif., May 31, 1967.
2. Harrje, D. T., "Fundamental Problems of Injector Design," presented at the 25th AGARD Colloquium on Advances in Tactical Rocket Propulsion, University of California, San Diego, Calif., Apr. 1965.
3. Conn, T. E., Hester, J. N., and Valentine, R. S., "Environmental Effects Upon Rocket Injector/Chamber Compatibility," *J. Spacecraft Rockets,* Vol. 4, No. 12, p. 1581, Dec. 1967.
4. Knight, R. M., Hawkins, D. E., and Rankel, R. E., *Chamber Technology for Space-Storable Propellants, 18th Quarterly Progress Report,* Contract NAS7-304, Report RM1390-340. Rocketdyne, Canoga Park, Calif., Apr. 21, 1969.
5. Dickerson, R., Tate, K., and Barsic, N., *Correlation of Spray Injector Parameters with Rocket Engine Performance,* Technical Report AFRPL-TR-68-147. Rocketdyne, Canoga Park, Calif., Jun. 1968.
6. Rupe, J. H., *On the Dynamic Characteristics of Free-Liquid Jets and a Partial Correlation with Orifice Geometry,* Technical Report 32-207. Jet Propulsion Laboratory, Pasadena, Calif., Jan. 15, 1962.
7. Carlson, L. W., *Spacecraft Rocket Engine Chamber Insulation Materials,* Report R-7838. Rocketdyne, Canoga Park, Calif., Apr. 30, 1969.

XXI. Systems Analysis

MISSION ANALYSIS DIVISION

A. Trajectory Considerations for an Earth to Jupiter to Saturn to Pluto Mission, *R. A. Wallace*

The 1975 to 1980 time period will afford an unusually favorable opportunity to launch missions to explore all of the outer planets. During this time period, the positions of the outer planets are such that multiple-planet trajectories involving all of these planets are possible (SPS 37-54, Vol. III, pp. 35–36). Some of these opportunities will not occur again for almost two centuries. One mission travels to all of the outer planets except Pluto (Ref. 1). This mission requires a close flyby of Saturn. Since the Rings of Saturn may present a formidable obstacle to missions with trajectories passing close to them, two alternative missions have been suggested. One has a trajectory which travels directly to Uranus and Neptune after Jupiter, bypassing Saturn. The trajectory considerations for this mission have been discussed in SPS 37-55, Vol. III, pp. 17–27. This article focuses on a second alternative mission; one that travels first to Saturn after Jupiter and then on to Pluto.

The Jupiter launch opportunities were searched for trajectories which proceeded on to Saturn and Pluto. Before 1976, the planets are aligned in such a way that too great a bending angle is required at Jupiter to proceed to Saturn, i.e., negative Jupiter flyby altitudes would be required. After 1978, Jupiter has rotated too far in its orbit to provide significant gravity assist for an Earth to Jupiter to Saturn to Pluto (E–J–S–P) mission. Thus, good launch opportunities occur in 1976, 1977, and 1978.

A possible problem for an E–J–S–P mission might be caused by a low value of the geocentric declination magnitude at Saturn and Pluto encounters. The geocentric declination of a spacecraft in flight is important to the accuracy of the orbit determination. The declinations at Jupiter encounter are 10 deg or more. Declinations for Saturn and Pluto encounters are 4 deg or less. Values less than 5 deg require investigation.

To provide an overall view of the important trajectory design factors, design charts were constructed. For outer-planet missions where flight times are long, a launch date–mission duration grid was found to be most effective. Figure 1 shows the launch energy C_3 contours on the launch date–mission duration grid. The best 30-day launch period was chosen with regard to C_3 for each

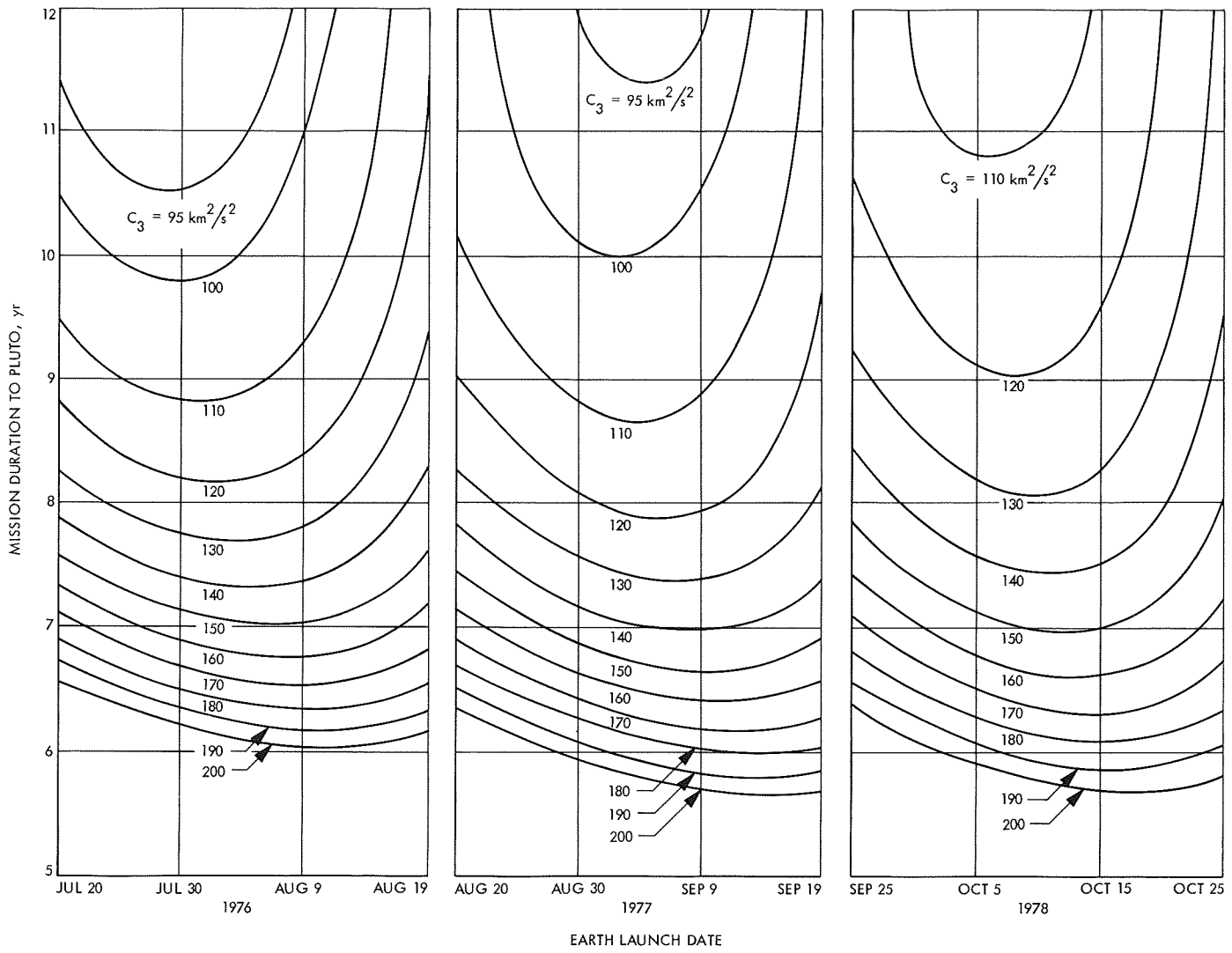


Fig. 1. E–J–S–P launch energy C_3 requirements

launch year. Only missions with durations below 12 yr and launch energies of 200 km^2/s^2 or below are considered. From a launch energy–mission duration standpoint, the 1977 launch opportunity is clearly superior to the other two.

A composite plot (using Fig. 1 as a base) was constructed considering the following constraints:

(1) Launch energies below 200 km^2/s^2.

(2) Mission durations to Pluto less than 12 yr.

(3) Positive flyby altitudes at Jupiter and Saturn.

(4) Encounter at Jupiter, Saturn, and Pluto at times other than a 20-day period before or after solar conjunction (to prevent communications blackout).

(5) The absolute value of the geocentric declination of the launch asymptote less than 28.5 deg (to prevent range safety and launch window problems).

Those trajectories allowable according to the above criteria are shown in Fig. 2. The shaded areas violate one of the restrictions stated.

Since the 1977 launch opportunity is superior to the other two, two missions were selected for launch in this year. One mission requires a maximum launch energy of 110 km^2/s^2 and a 15-day launch period. The other mission requires a larger maximum launch energy of 120 km^2/s^2, but affords a longer launch period of 21 days. A factor that determines which mission is more attractive will certainly be the availability and capability of the

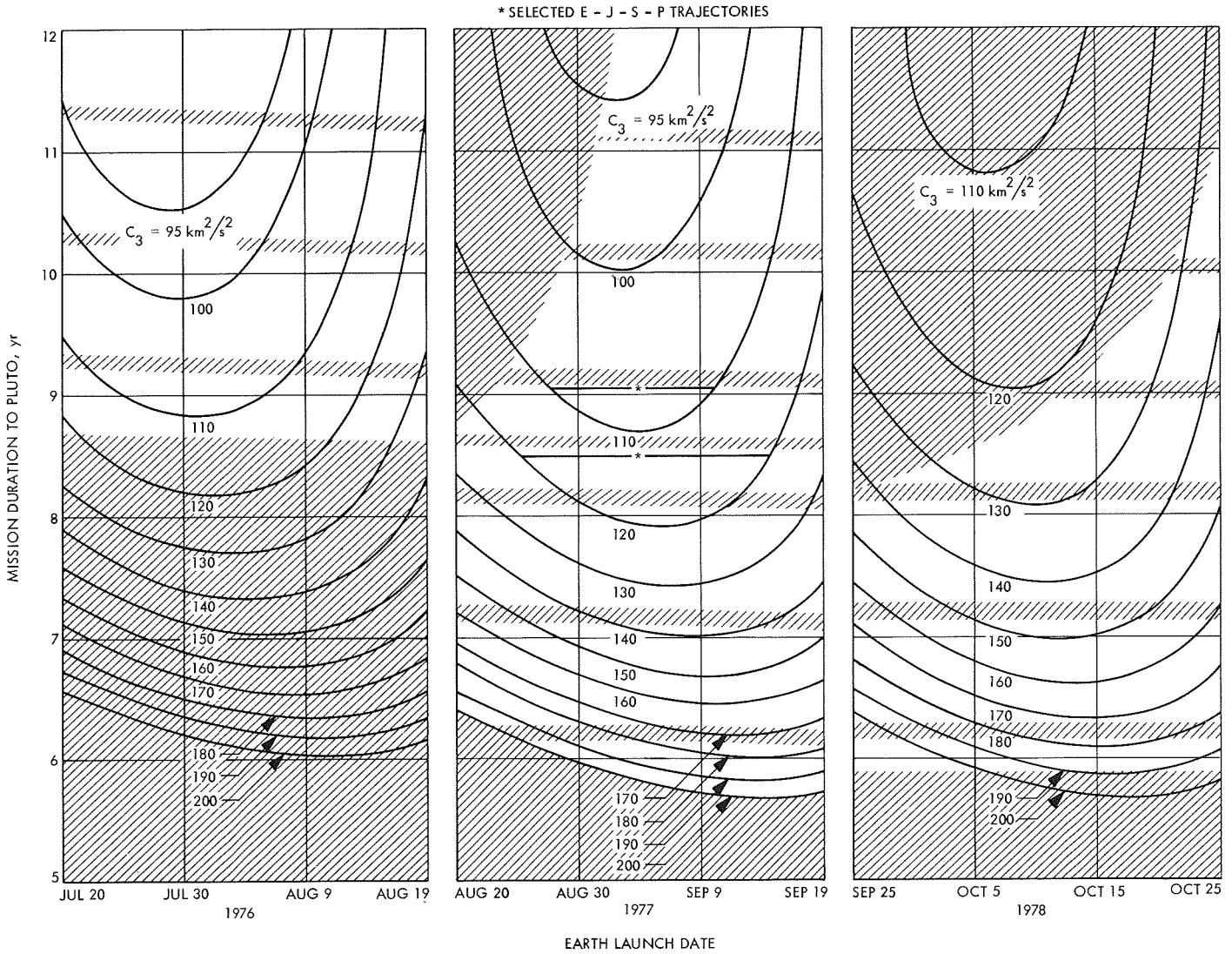


Fig. 2. E–J–S–P candidate trajectories

launch vehicle used. A trajectory from the middle of each launch period was selected to present typical trajectory characteristics. The planet encounter dates are given for each trajectory in Table 1.

Table 1. Mission characteristics

Characteristic	110-km^2/s^2 mission (max launch energy)	120-km^2/s^2 mission (max launch energy)
Launch energy, km^2/s^2	107.0	112.1
Earth launch date	Sep. 4, 1977	Sep. 4, 1977
Jupiter arrival date	Mar. 1, 1979	Feb. 6, 1979
Saturn arrival date	Nov. 7, 1980	Sep. 12, 1980
Pluto arrival date	Sep. 21, 1986	Mar. 9, 1986
Mission duration, yr	9.1	8.5

The 1977 E–J–S–P trajectories and launch periods are indicated in Fig. 2, and a heliocentric ecliptic plane plot is shown in Fig. 3. The major difference between the two trajectories is the greater bending required at Jupiter and Saturn by the 120 km^2/s^2 launch energy mission. The combination of closer flybys and greater launch energy results in a 6-mo reduction in duration for this mission.

Flyby trajectories are principally influenced by the closest approach altitude and approach velocity which determine the amount of trajectory bending during encounter. Flyby trajectory characteristics are compared in Table 2 for both trajectories.

The effects of flyby altitude and approach speed can most easily be seen in the heliocentric plot and trajectory

Table 2. Flyby trajectory characteristics

Characteristic	110-km^2/s^2 mission (max launch energy)	120-km^2/s^2 mission (max launch energy)
Jupiter flyby altitude, km	248,000 (3.5 R_J)[a]	230,000 (3.2 R_J)[a]
Jupiter approach speed, km/s	11.1	11.8
Jupiter flyby inclination, deg	4.0	3.9
Saturn flyby altitude, km	561,000 (9.3R_S)[b]	452,000 (7.5 R_S)[b]
Saturn approach speed, km/s	15.4	16.3
Saturn flyby inclination, deg	81.3	71.3
Pluto flyby altitude, km	Open	Open
Pluto approach speed, km/s	16.9	18.5

[a] R_J = Jupiter surface radius.
[b] R_S = Saturn surface radius.

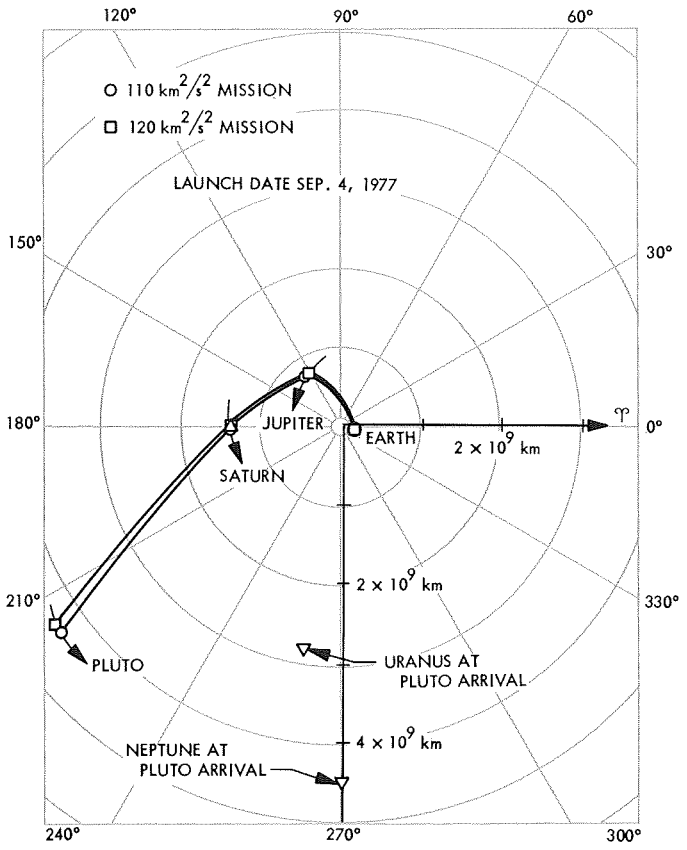


Fig. 3. Heliocentric ecliptic plane plot (1977 E–J–S–P)

encounter profiles (Figs. 4 and 5). Since both trajectories are qualitatively similar, only one is shown. Note in Table 2 the high flyby inclination required at Saturn. The aiming point is near the south pole and thus results in the spacecraft passing underneath the planet. This high inclination is required because the orbit of Pluto is inclined to the ecliptic by some 17 deg and is near maximum heliocentric latitude at arrival. Since no Pluto post-encounter mission has been considered yet, the choice of aiming point is open. To illustrate the small effect Pluto might have on a close encounter, a trajectory profile plot is shown in Fig. 6 for a closest approach altitude of one Pluto radius (7016 km).

It is of interest to know how the target planet would look if viewed from the spacecraft at various times during encounter. Figures 7 and 8 show Jupiter and Saturn as viewed from the spacecraft throughout the encounter. Again only one trajectory was used, since the other is similar. Projections of the vectors to the sun and earth on the plane perpendicular to the planet–spacecraft vector are also shown.

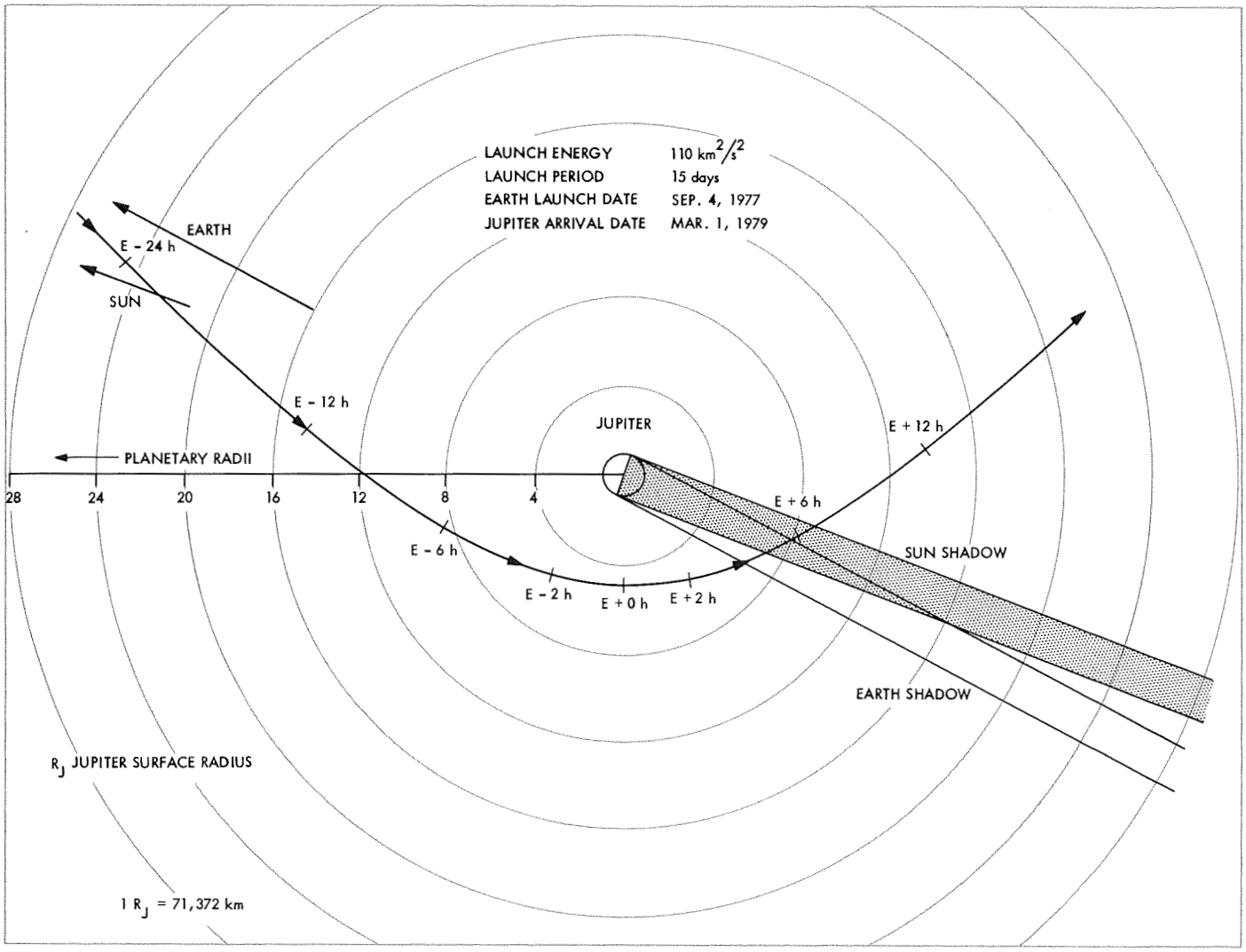


Fig. 4. Jupiter encounter (1977 E–J–S–P)

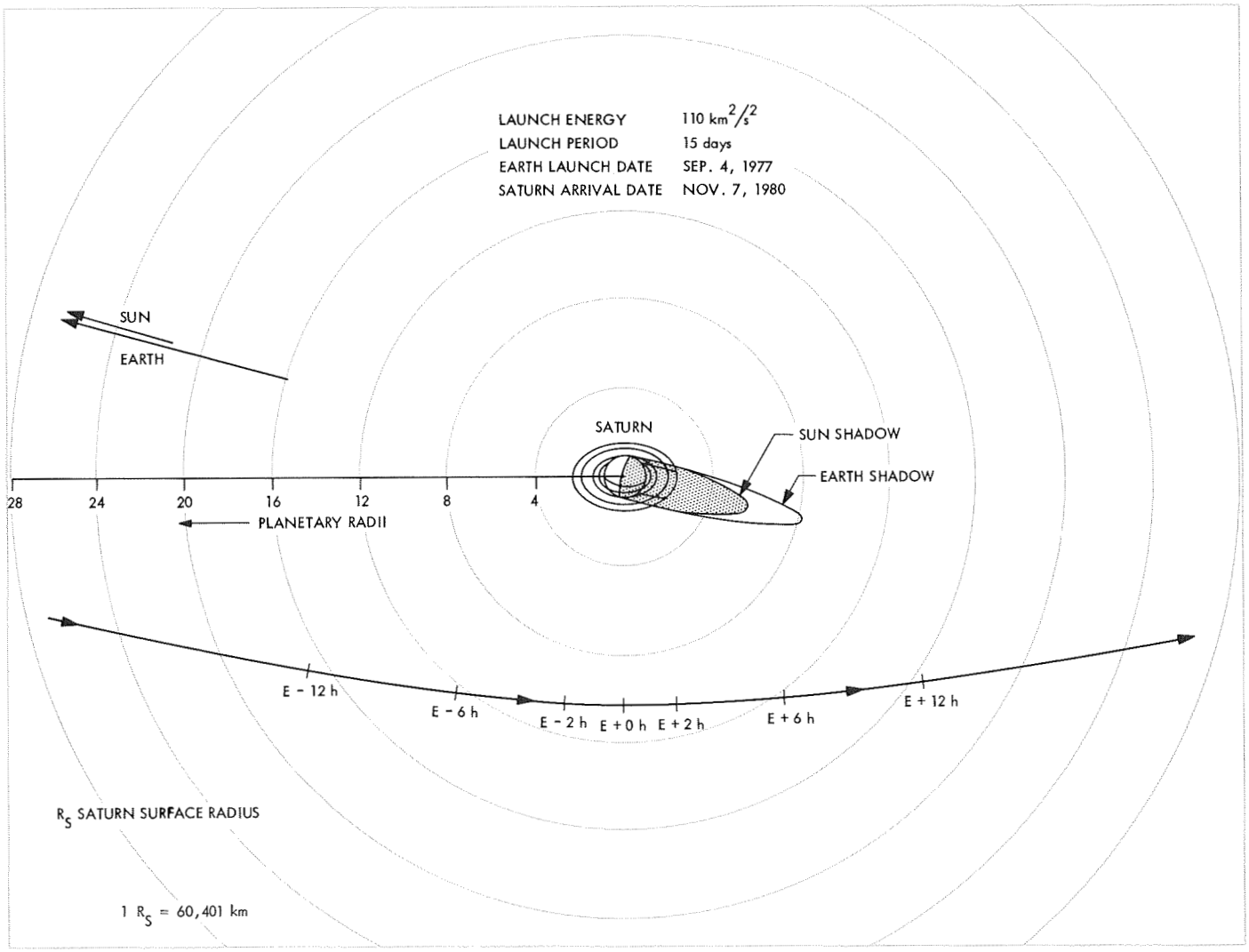


Fig. 5. Saturn encounter (1977 E–J–S–P)

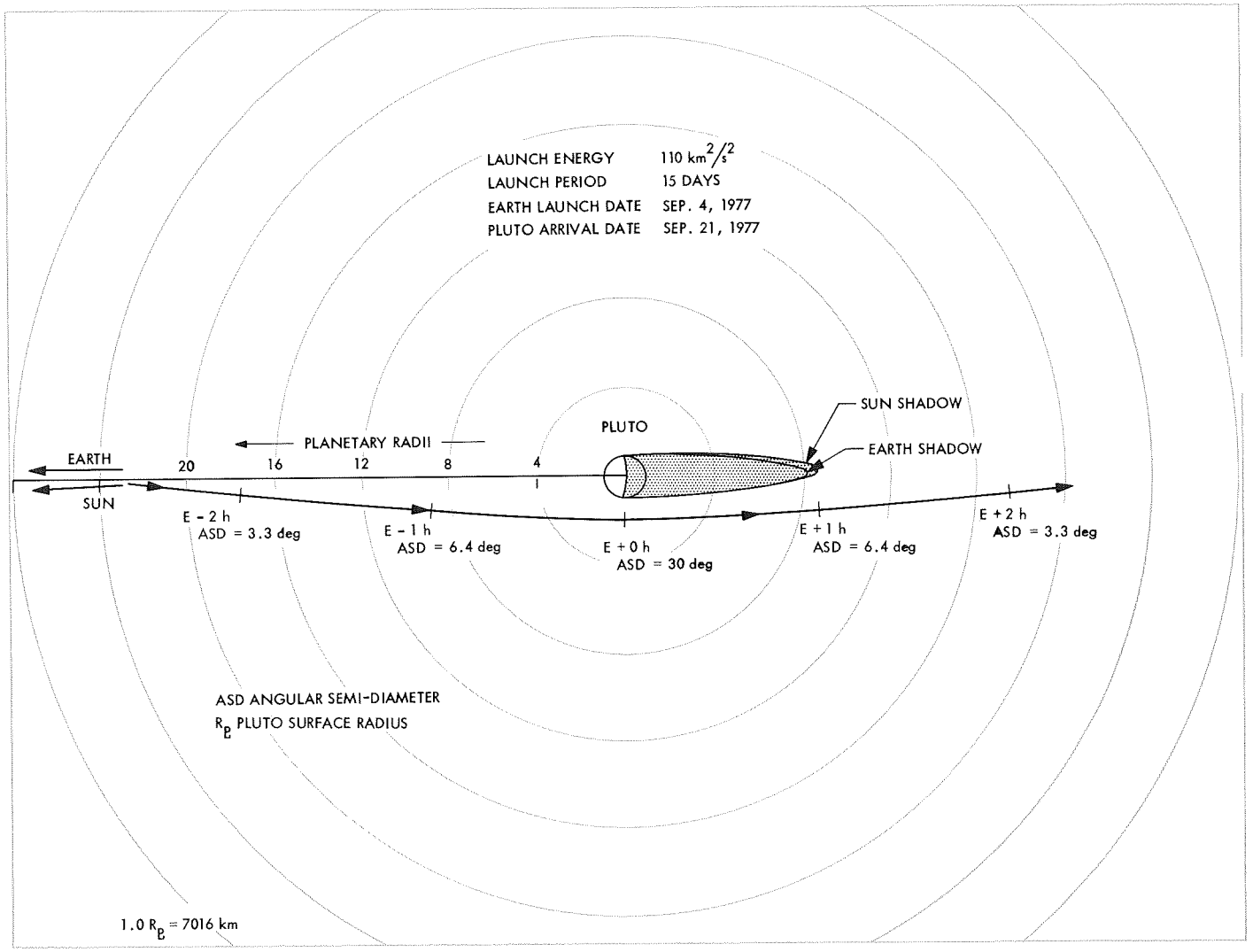


Fig. 6. Pluto encounter (1977 E–J–S–P)

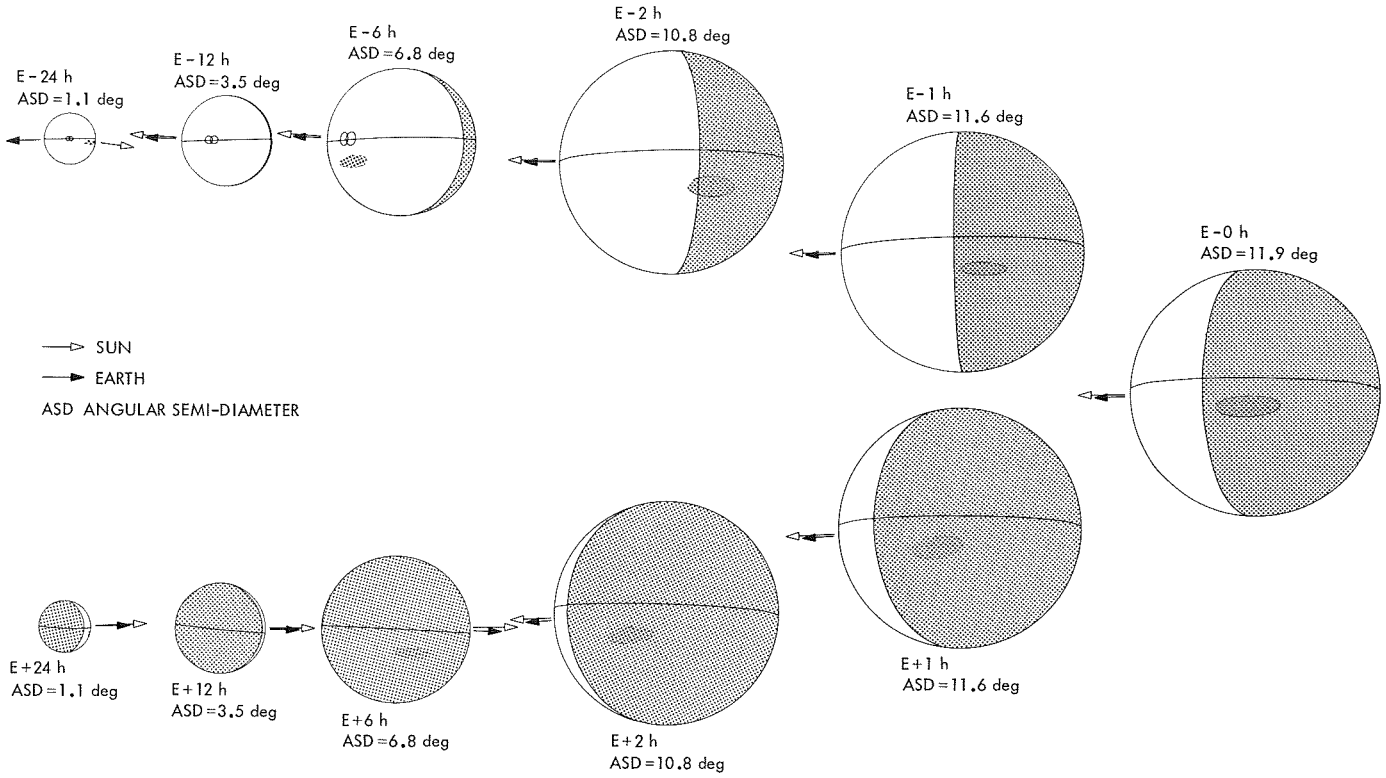


Fig. 7. Views of Jupiter from spacecraft (1977 E–J–S–P, launch energy = 110 km²/s, launch period = 15 days)

Due to multiple-planet trajectory requirements, aiming points at Jupiter and Saturn are fixed; the aiming point at Pluto is open for selection for the optimum flyby science or post-encounter exploration. For the two trajectories selected, earth occultation does not result at Saturn and flyby altitudes are fairly large at that planet. This is a result of the position of Pluto in space at arrival there (Fig. 3).

The principal factors affecting the design of a grand tour mission will also affect the design of an E–J–S–P mission, with two exceptions. The most important exception is that the Saturn encounter for the E–J–S–P mission has a large closest approach altitude, and therefore the Rings will not pose an impact problem. This advantage is offset somewhat by the degradation of possible Saturn science experiments; earth occultation is not possible and for the trajectories selected Saturn will never have an angular semi-diameter larger than 7 deg. Also, the trade-off for reaching Pluto rather than having a second flyby (or possibly first) of Uranus and Neptune must be considered. A possible problem for orbit determination accuracy has also been identified with the finding that Saturn and Pluto geocentric declinations at encounter are less than 4 deg. Declinations less than 5 deg have resulted in accuracy problems in the past, and further investigation would be in order. Major design factors for the E–J–S–P mission remain the same as for the grand tour mission:

(1) Mission durations in excess of 8 yr.

(2) Launch energies over 100 km²/s².

(3) Communication distances close to 30 AU.

(4) Round-trip signal transmission times of up to 8 h.

(5) Precision trajectory guidance.

Reference

1. Kingsland, L., Jr., "Trajectory Analysis of a Grand Tour Mission to the Outer Planets", AIAA Paper 68-1055, Philadelphia, Pa., Oct. 1968.

B. A Preliminary Study of a Solar-Electric Low-Thrust Saturn Orbiter Mission in 1983–84,

D. J. Alderson and W. Stavro

1. Introduction

Recently, solar-electric, low-thrust missions to the outer planets have received a great deal of attention. An intensive effort is being undertaken to explore the outer planets

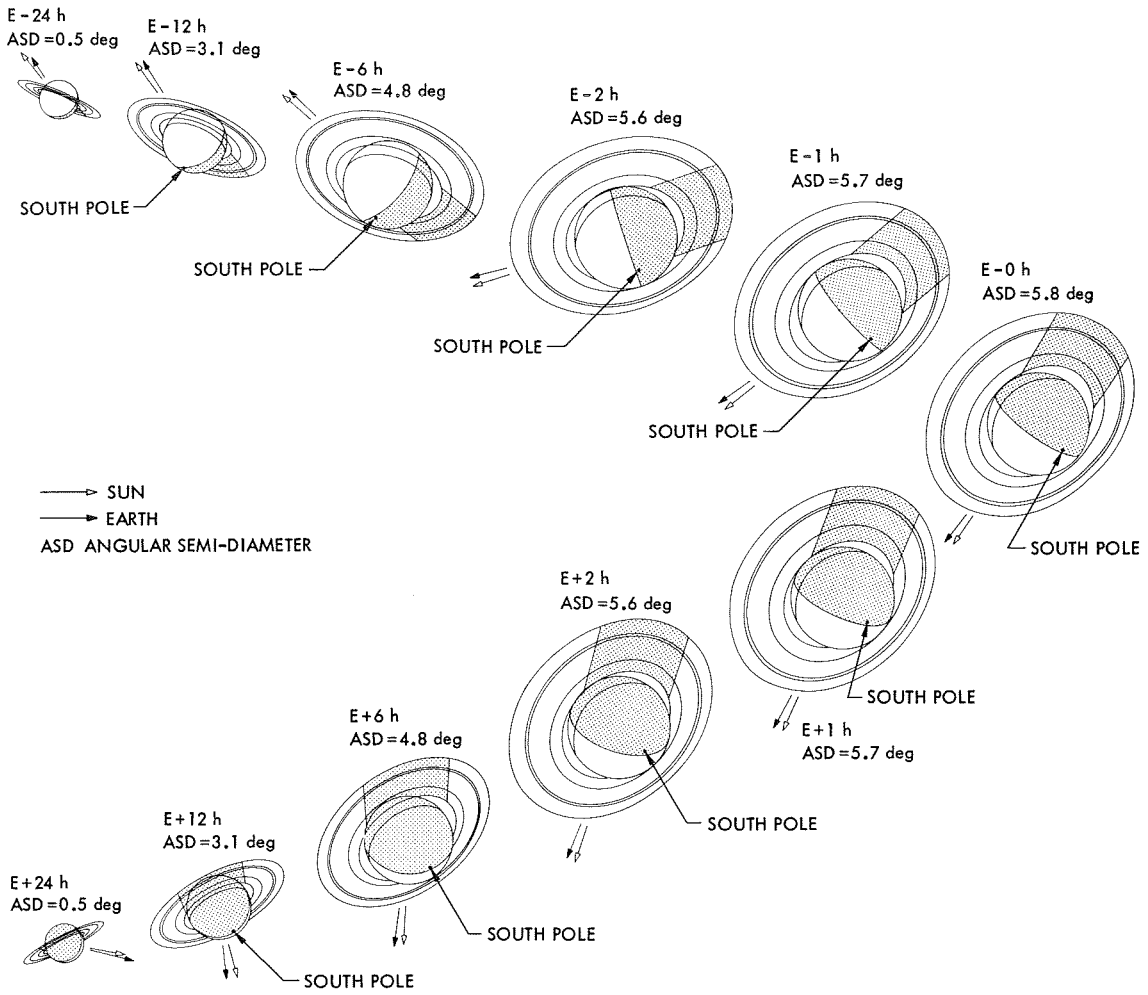


Fig. 8. Views of Saturn from spacecraft (1977 E–J–S–P, launch energy = 110 km²/s, launch period = 15 days)

in the next two decades with heavy emphasis placed on the grand-tour type missions in the late 1970s. These missions use gravity-assist trajectories and are, therefore, all flybys. Studies are in progress to investigate orbiter missions to the outer planets in the late 1970s and early 1980s using ballistic trajectories. Such studies have shown that performance and payloads suffer greatly for planets past Jupiter. It has been shown[1] that a ballistic orbiter mission to Saturn in the early 1980s will have trouble inserting a 1000-lb payload in an orbit around Saturn that would stay outside the rings, using a *Titan III* × (1207)/*Centaur* launch vehicle with a *Burner II*. Such a mission is only possible with flight times in excess of 4.3 yr.[2]

Therefore, it was decided to investigate a mission similar to the most optimum ballistic case (1984 opportunity), using low-thrust electric propulsion. The attempt was to determine some trajectory and performance parameters in order to compare them to the ballistic case. Since a detailed mission definition was not the objective, a level-1 computer program was used for this study. (A level-1 program is one used in generating basic trajectory performance parameters, using approximate methods to reduce running time.) In this analysis, it was decided to use the level-1 program ASTRAL, which is an optimized low-thrust trajectory program that uses approximate closed-form solutions to the equations of motion.

2. ASTRAL

ASTRAL was described in SPS 37-54, Vol. III, pp. 20–26. Since the writing of that article, the standard trajectory model based on the analysis of P. Wesseling (Ref. 1)

[1]Stavro, W., *Saturn Orbiter Possibilities for the 1980–85 Launch Opportunity* (JPL internal document).

[2]Stavro, W., *1983–84 Saturn Orbiter Opportunity* (JPL internal document).

has been modified only slightly. Since the most severe deviations of the closed-form approximations from exact solutions involved the time as a function of angular position in the ecliptic, this variable was changed to a numerical integration of the closed-form expression for solar distance, resulting in a substantial increase in accuracy.

Two alternative trajectory models have been added to ASTRAL. The first involves direct numerical integration of the equations of motion and, thus, is very useful for purposes of comparison and evaluation of individual trajectories, but executes too slowly on the computer to be useful in trajectory optimization in the ASTRAL context. The second alternate trajectory model was constructed in an attempt to overcome the limitations imposed by Wesseling's assumption that the eccentricity of the spacecraft orbit is of the same order as ϵ, the ratio of spacecraft thrust acceleration to the acceleration of local solar gravity. Unfortunately, the removal of this restrictive assumption led to such additional complications in the derivation of the closed-form approximation that it could only be carried out to a low order in ϵ. As a result, the hoped-for improvement in accuracy for large orbital eccentricities did not result. Therefore, the standard trajectory model was used in this study. Thrust ratios ϵ are all near 0.1 for our optimized trajectories, while the orbital eccentricities at the start of the low-thrust propulsion phase have been in the range of 0.4 to 0.6.

Three forms of optimizers have been tried with ASTRAL, and the two most successful retained as options. The third, a gradient projection method, involved assumptions that were not appropriate in the ASTRAL environment. One of the retained optimizing methods is M. J. Box's complex method (SPS 37-54, Vol. III). The other method, which produced better results than Box's complex method in many cases, is now the standard ASTRAL optimizer. It is a direct-search method like the complex method, but operates on the basis of a single point in the parameter space rather than a sizable collection, or complex, of such points. The one-point search routine uses pseudo-random numbers to select a set of orthogonal test vectors that span the parameter space. It then evaluates the function to be maximized at points formed by adding and subtracting the test vectors from the reference vector and tries to move the reference vector in those directions in which the function is increasing. No comparisons are made of the rates of increase in different directions in the parameter space, thus avoiding various difficulties with scaling and non-differentiable functions that plague gradient methods. If no increase is achieved in the function value, the size of the test vectors is halved and, if they are still larger than a specified limit, the search process is repeated. Both one-point and complex optimizers were used in this study.

3. Mission Assumptions

The study used two-dimensional trajectories throughout, with no plane-changing maneuvers included. Saturn's orbit, inclined by 2.5 deg to the ecliptic, was approximated by its projection on the ecliptic, and the final orbit of the spacecraft around Saturn was assumed to lie in the ecliptic.

The mission flight pattern consisted of impulsive launch from earth orbit, followed by a single low-thrust flight phase, subsequent unpowered coast to the vicinity of Saturn, and final impulsive transfer into orbit around Saturn. During the low-thrust stage, the spacecraft thrust vector is maintained in constant orientation to the spacecraft–sun line in the ecliptic plane, and the magnitude of the spacecraft thrust acceleration is taken to be inversely proportional to the square of the solar distance.

The launch vehicle used was the *Titan III* × (1207)/*Centaur* without a *Burner II*. Its payload capability was approximated by fitting a simple exponential function to more precise payload data in the C_3 range from 0 to about 80 km^2/s^2. The resulting function, as shown in Fig. 9, was

$$M_0 = A_1 \exp(-\Delta V_1/C_1) - B_1$$

$$\Delta V_1 = (V_{e1}^2 + C_3)^{1/2} - V_{e1}$$

$$A_1 = 9064 \text{ kg}$$

$$B_1 = 1424 \text{ kg}$$

$$C_1 = 3.324 \text{ km/s}$$

$$V_{e1} = 11.021 \text{ km/s}$$

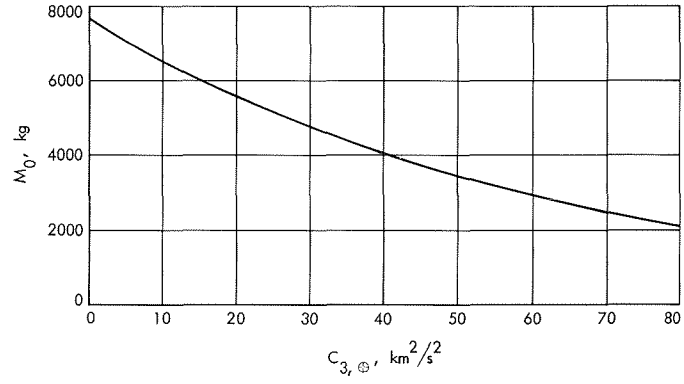


Fig. 9. *Titan III* × (1207)/*Centaur* performance curve

where M_0 is the spacecraft mass injected into hyperbolic earth orbit with specified C_3, ΔV_1 is the velocity change required in excess of that needed to achieve parabolic orbit, and V_{e1} is the escape velocity at the altitude of the initial earth orbit from which the spacecraft is launched (100 nmi).

The low-thrust propulsion system was taken as having a specific impulse of 3500 s, a specific powerplant weight of 34 kg/kW, and a conversion efficiency from electric to kinetic energy of 69.7%. The orbit insertion engine used for impulsive deboost at Saturn was taken as having a structure-to-total-weight ratio of 0.20 and a specific impulse of 385 s.

The deboost maneuver at Saturn was performed at 2.3 Saturn radii, or 139,000 km, from the center of the planet, which thus became the periapsis distance for both the approach hyperbola and the final orbit. The latter was given a period of 30 days, and thus a semimajor axis of 1,862,000 km and an eccentricity of 0.925.

With this orbit insertion engine and target orbit, the result, as shown in Fig. 10, is

$$M_{PL} = M_F (A_2 \exp(-\Delta V_2/C_2) - B_2)$$

$$\Delta V_2 = (V_{e2}^2 + V_\infty^2)^{1/2} - V_P$$

$$A_2 = 1.25$$

$$B_2 = 0.25$$

$$C_2 = 3.77556 \text{ km/s}$$

$$V_{e2} = 23.367 \text{ km/s}$$

$$V_P = 22.927 \text{ km/s}$$

where M_{PL} is the spacecraft usable payload mass in the final orbit, M_F is the mass of the payload plus orbit insertion engine and fuel just before the deboost maneuver, ΔV_2 is the velocity change produced by the engine, V_{e2} is the Saturnian escape velocity at the maneuver distance, V_∞ is the hyperbolic excess velocity of the spacecraft in its hyperbolic approach trajectory, and V_P is its periapsis velocity in its final orbit.

Parameters that are optimized in this study are

(1) Direction and magnitude of hyperbolic excess velocity vector at launch.

(2) Spacecraft low-thrust power at 1 AU.

(3) Direction of the low-thrust vector.

(4) Point at which low-thrust power is cut off.

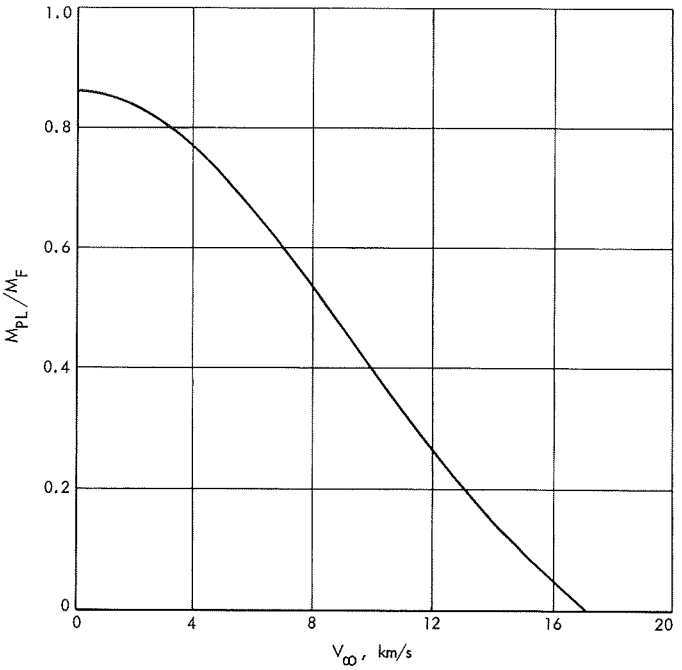


Fig. 10. Orbit insertion motor performance curve

4. Discussion of Results

The first attempt in the study of this mission was to determine the optimum launch dates. Figure 11 shows a plot of launch date versus payload. The flight time is fixed at 1280 days. It is noted from Fig. 11 that for a launch period of 20 days a payload of 585 kg may be obtained; and for a launch period of 15 days, a 615-kg payload may be achieved. Figure 12 shows the heliocentric trajectory for the December 19, 1983 launch date.

Figure 11, showing payload versus launch date, was obtained by making six individual optimized runs at launch dates that differ by 5 days. Some of the parameters of these runs are shown in Table 3.

Figure 13 shows the orbit geometry at Saturn. Saturn's rings and the orbits of its satellites are also shown for comparison. The satellite orbits are shown as rotated into the ecliptic plane.

5. Conclusions

The above analysis is a preliminary study of a solar-electric orbiter mission to Saturn in the 1983–84 launch opportunity. The analysis was performed using a level-1 low-thrust computer program (ASTRAL), and the accuracy of the results are, therefore, limited by the accuracy of the program, which is discussed in SPS 37-54, Vol. III.

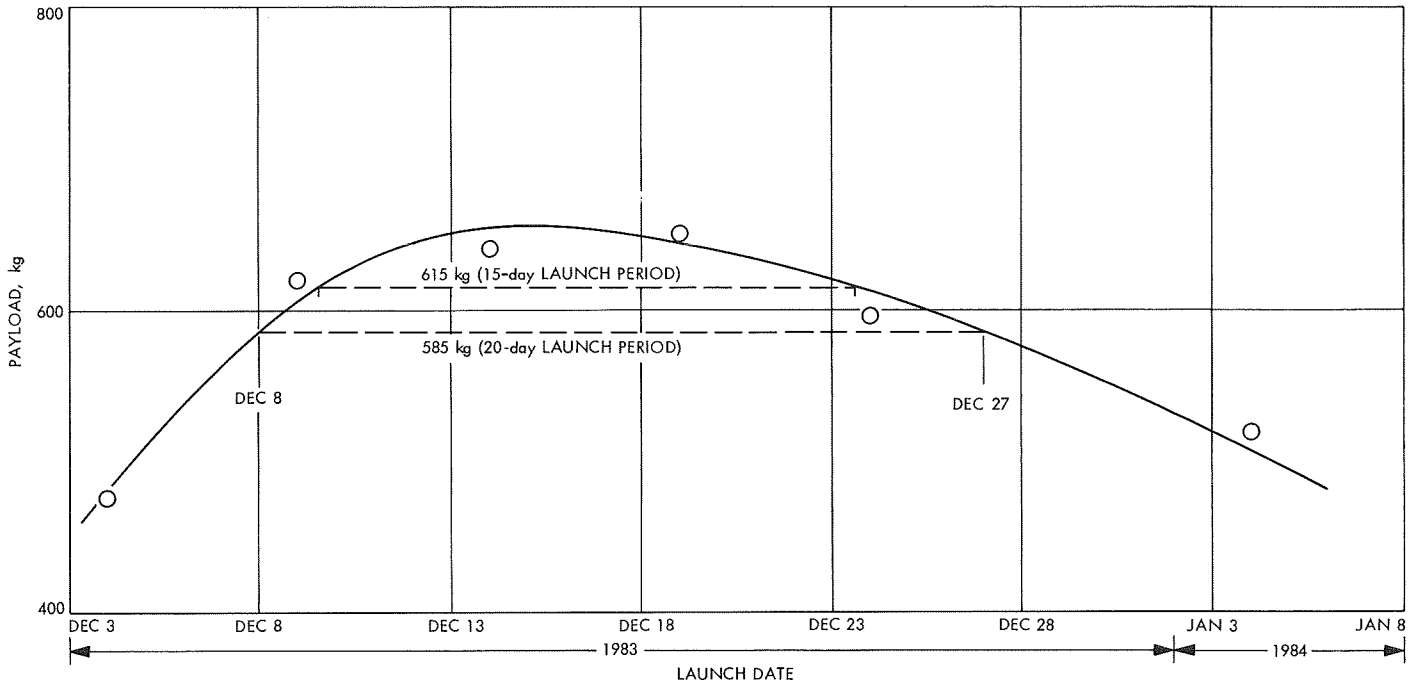


Fig. 11. Payload vs launch date curve for fixed flight time

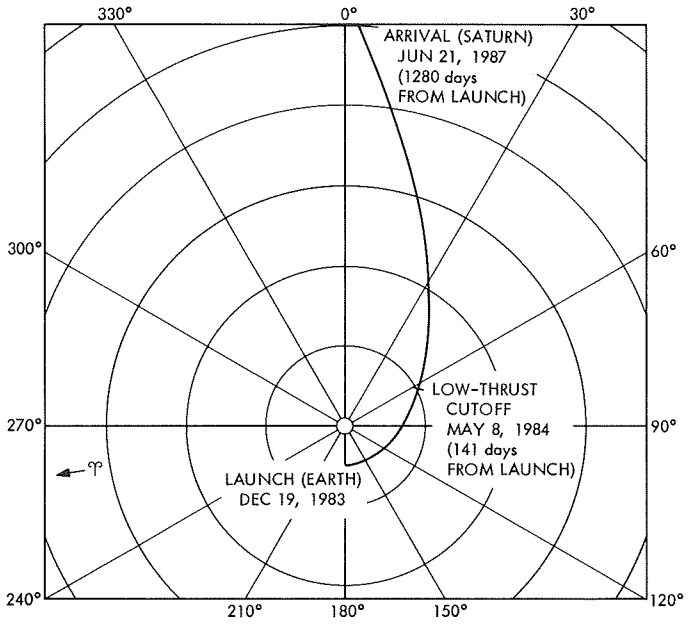


Fig. 12. Heliocentric trajectory of a solar-electric orbiter to Saturn

The study indicates significant advantages of using solar-electric low-thrust propulsion for such a mission. Flight times were maintained at 3½ yr, and with insertion beyond the rings of Saturn much higher payloads were achieved than in the ballistic case.

Reference

1. Wesseling, P., "A Two-Variable Asymptotic Solution for Three-Dimensional Solar-Powered Low-Thrust Trajectories in the Vicinity of the Ecliptic Plane," *Astronaut. Acta*, Vol. 13, pp. 43–440, 1967.

Table 3. Parameters of optimized runs

T_L	ϕ_p, deg	ϕ_c, deg	M_{PL}, kg	M_P, kg	$d_{\oplus-sc,A}$, km × 10^6	$d_{\odot-sc,A}$, km × 10^6	$C_{3,L}$, km^2/s	$V_{\infty,A}$, km/s	ΔV_A, km/s	Pow_{1AU}, kW
Dec. 4, 1983 (JD 2445673)	113.5	72.7	477	573	1351	1503	41.4	9.64	2.35	65.9
Dec. 9, 1983 (JD 2445678)	128.6	52.7	619	568	1351	1503	39.4	9.66	2.36	59.7
Dec. 14, 1983 (JD 2445683)	116.3	60.0	640	494	1352	1503	41.8	9.58	2.33	56.4
Dec. 19, 1983 (JD 2445688)	109.8	61.6	649	385	1354	1503	49.3	9.53	2.31	46.3
Dec. 24, 1983 (JD 2445693)	130.2	36.3	596	515	1357	1503	45.0	9.56	2.32	53.2
Jan. 3, 1984 (JD 2445703)	116.2	40.4	520	362	1367	1503	59.5	9.43	2.27	40.7

T_L = time of launch; ϕ_p = length of powered flight arc; ϕ_c = length of coast arc; M_{PL} = mass of payload; M_P = mass of low-thrust propellant; $d_{\oplus-sc,A}$ = earth–spacecraft (communication) distance at arrival; $d_{\odot-sc,A}$ = sun–spacecraft distance at arrival; $C_{3,L}$ = injection energy at launch; $V_{\infty,A}$ = hyperbolic excess velocity at arrival; ΔV_A = velocity increment for orbit insertion at arrival; Pow_{1AU} = spacecraft low-thrust electric power at 1 AU from the sun.

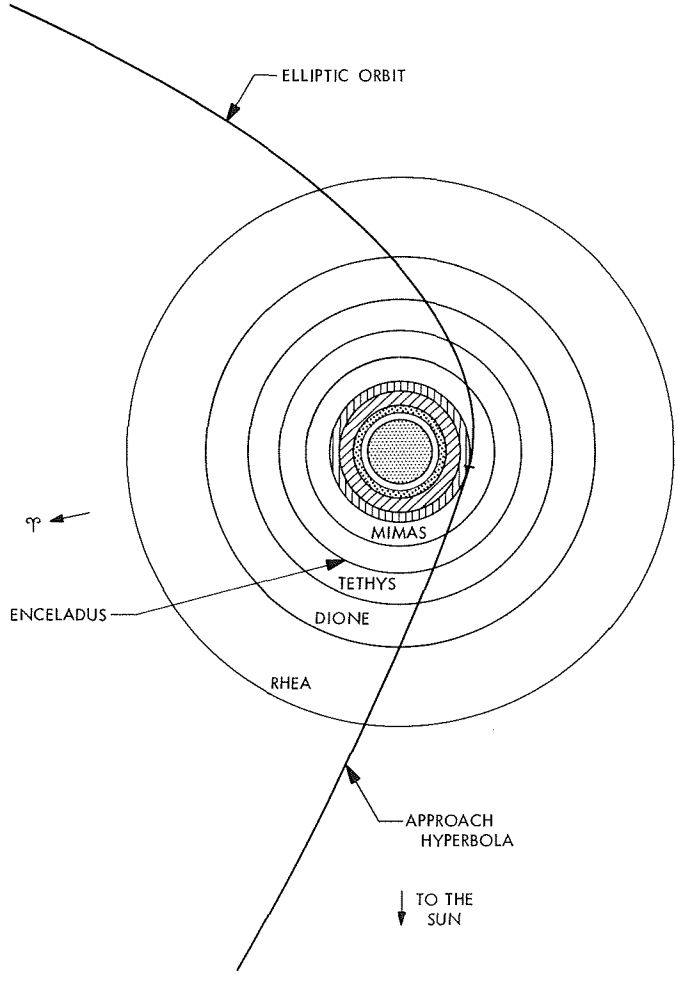


Fig. 13. Orbit geometry showing Saturn's rings and satellites

C. On the Optimal Number of Remote Computer Terminals, *R. G. Chamberlain*

1. Introduction

Obtaining more remote computer terminals than are absolutely necessary to handle an organization's computing load would lead to idle terminals. On the other hand, obtaining only the absolutely necessary number of terminals would, as a result of queuing, lead to inconvenience and delays. Figure 14 shows the results of a small study to determine the optimal number of terminals under various conditions.

Figure 14 shows the number of terminals which minimizes the sum of the cost of idle terminal time and the cost of delays to engineers or programmers. Implicit in Fig. 14 is the assumption that the costs of delays are linearly related to the lengths of such delays. It is further assumed that these are real costs, even though they generally show up as decreased productivity, rather than as money out of pocket.

The queuing problem was approached in two ways: by simulation and by analysis. The simulation utilized data gathered from a dozen engineers about expected usage of terminals. It took into account estimated probability distributions of the times between needs for a terminal, times of day of such needs, and terminal usage times. These results are shown as crosses on Fig. 14. The analysis embodied the simplifying assumptions that inter-arrival and service times are exponentially distributed and that all times of the working day (8 a.m. to 5 p.m.) are equally likely to be used for computation. Both approaches agree surprisingly well. The lower edge of each band in Fig. 14 represents the assumption that each engineer contributes 0.02 (that is, 3.8 h of terminal time per month) to ρ. The upper edge results from assuming a contribution of 0.05 (9 h per month).

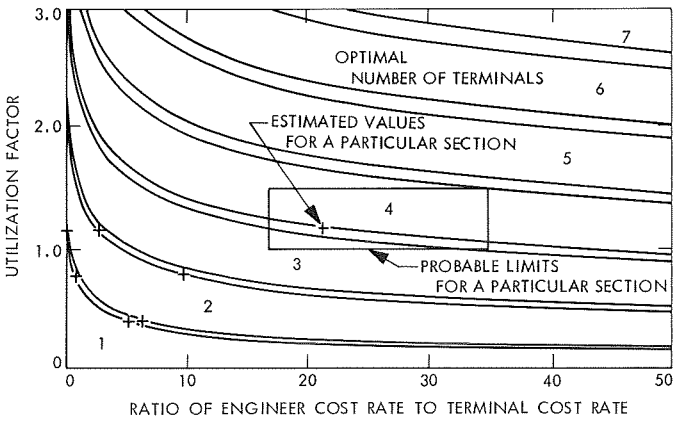


Fig. 14. Optimal number of terminals

If more terminals are obtained than are absolutely necessary to service the computing needs, there will be times when one or more terminals are idle. These idle periods represent a cost, as shown by the small circles in Fig. 15. If use of the terminals could be completely scheduled with no loss of efficiency or effectiveness, this (plus some figure which does not vary with the number of terminals) would be the total cost. This total cost would be minimized by the smallest number of terminals which could do the job.

If, on the other hand, the need for terminals cannot be completely scheduled without loss of efficiency, there will occasionally be more customers than terminals, resulting in more or less costly delays. The larger the number of terminals, the lower the costs of such delays. This cost is shown in Fig. 15 by small squares. It is not linear with the number of terminals, since it depends upon the interaction of probabilities and the cost versus delay time functions.

The total (relevant) cost is the sum of these two costs, and is shown in Fig. 15 by small crosses. Since one cost component increases with the number of terminals, and the other cost component decreases, there is some number

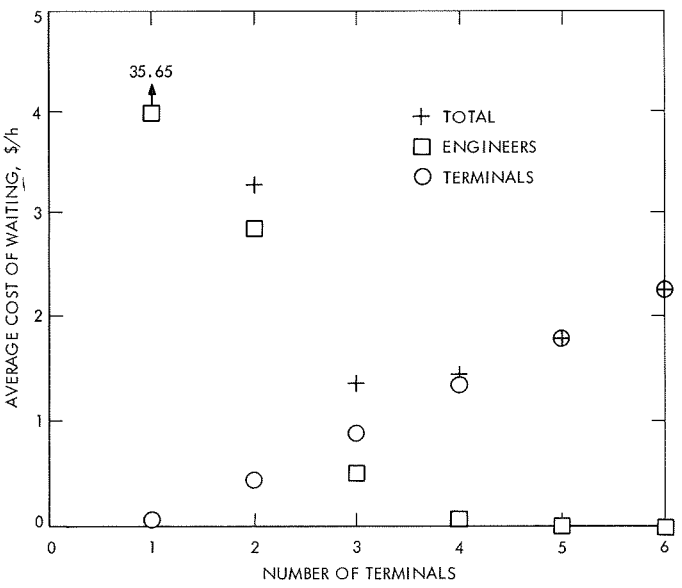


Fig. 15. Dependence of waiting costs on number of terminals

of terminals that gives the minimum total cost. The purpose of this analysis is to determine that optimal number of terminals.

One way to achieve this purpose would be to prepare Fig. 15 for a variety of terminal usage factors and a variety of delay cost functions. Fortunately, however, if we are willing to accept the assumption that delay costs are linear with delay time, the delay cost rate and the usage factor can be presented parametrically.

2. Parameterization of Results

In order to optimize the number of terminals, we need not consider all the costs involved in computing. It is sufficient to consider only those costs which are affected by the number of terminals. Thus, the total *relevant* cost is the sum of the cost of unused terminal time and the cost of delays:

$$TRC_i = TCR * TWT_i + ECR * EWT_i \qquad (1)$$

where TRC_i = total relevant cost incurred in some time (ST, say) when there are i terminals; TCR, ECR = terminal and engineering waiting cost rates, respectively; and TWT_i, EWT_i = terminal and engineer waiting time during ST, respectively, when there are i terminals.

Equation (1) can be normalized and made dimensionless by dividing by ST and TCR, giving Eq. (2).

$$y_i = A_i + B_i x \qquad (2)$$

where $y_i = TRC_i/(ST * TCR)$ = total relevant cost per unit time when there are i terminals, in units of the terminal cost rate; $x = ECR/TCR$ = ratio of engineer waiting cost rate to terminal waiting cost rate; $A_i = TWT_i/ST$ = average number of hours per hour that terminal(s) are idle when there are i terminals; and $B_i = EWT_i/ST$ = average number of hours of engineers' delay per hour when there are i terminals. If we know the values of A_i and B_i (either by solving a simplified model, by simulation, or by other means), then we can prepare Fig. 16.

From Fig. 16, it is evident that we can find the optimal number of terminals even if we are uncertain what the cost ratio x really is. In particular, if $0 \leqq x < x_1$, then the optimal number of terminals is 1; if $x_1 < x < x_2$, it is 2; and, in general, if $x_{N-1} < x < x_N$, then the optimal number of terminals is N. These critical cost ratios, the x_i, are

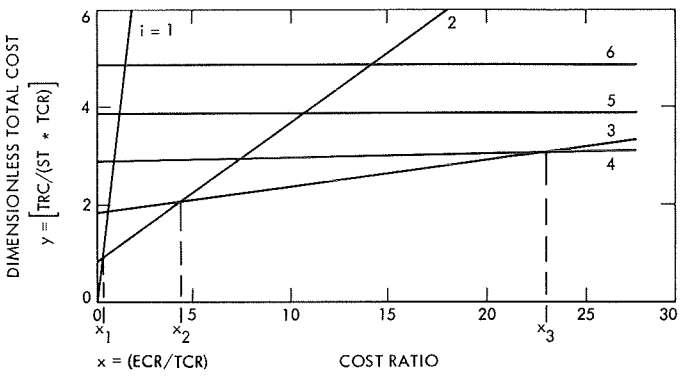


Fig. 16. Total cost dependence on engineer/terminal waiting cost ratio

simply the points where the lines of Eq. (2) cross. That is, they satisfy Eq. (3).

$$A_i + B_i x_i = A_{i+1} + B_{i+1} x_i \qquad (3)$$

If we now define the usage ratio ρ to be the fraction of time that terminals are actually in use (e.g., if two terminals could be kept busy all the time, ρ would be 2), and note that $A_i = i - \rho$, then the solution of Eq. (3) is

$$x_i = \frac{1}{B_i - B_{i+1}} \qquad (4)$$

The total cost y in Fig. 16 may be of interest, but it is no longer needed to find the optimal number of terminals. The x_i are the decision points, and their variation with usage, as in Fig. 14, completes the parameterization.

3. Analysis

The usual analytical approach to queuing problems can be easily applied in this case if the arrival and service distributions are assumed to be particularly simple (exponential), and if all hours of the working day are assumed to be equally desired. To study the effects of these assumptions, a program was written to simulate terminal usage and distributions were gathered from a dozen engineers. As indicated in Fig. 14, the two approaches give essentially the same results.

For the analytical model, the three simplifying assumptions are: (1) the probability that an engineer, if not using a terminal at time t, will want to get on a terminal between t and $t + dt$ is $\lambda\, dt$, (2) the probability that an engineer, if using a terminal at time t, will finish between t and $t + dt$ is $\mu\, dt$, and (3) these probabilities are the same for all engineers.

If we let p_n = the probability that n engineers are either using a terminal or waiting to use a terminal, N_E = the number of engineers, and N_T = the number of terminals, then

$$B_{N_T} = \sum_{n=N_T+1}^{N_E} (n - N_T)\, p_n \tag{5}$$

and

$$\rho = N_E \frac{\lambda}{\mu} \tag{6}$$

The set of difference equations can be derived by the method usually used in queuing problems: the transition equations

$$\{\text{of the form } p_n(t + dt) = f[p_{n-1}(t), p_n(t), p_{n+1}(t); \lambda\, dt, \mu\, dt]\}$$

are written and recognized as difference-differential equations

$$\left\{\text{of the form } \frac{p_n(t + dt) - p_n(t)}{dt} = f[p_{n-1}(t), p_n(t), p_{n+1}(t); \lambda, \mu]\right\}$$

Transient behavior is ignored by setting all derivatives to zero. This derivation is omitted here because it is lengthy but straightforward. The result is the following set of difference equations, which are easily solved numerically for the p_j $(j = 0, \cdots, N_E)$:

$$n = 0\text{: } p_1 = \rho p_0$$

$$0 < n < N_T\text{: } p_{n+1} = \left[\frac{\rho}{n+1} + \frac{n(1-\gamma)}{n+1}\right] p_n - \left(\frac{\rho}{n+1} - \gamma\right) p_{n-1}$$

$$N_T \leq n < N_E\text{: } p_{n+1} = \left(\frac{\rho}{N_T} + 1 - \frac{n\gamma}{N_T}\right) p_n - \left[\frac{\rho}{N_T} - \frac{(n-1)\gamma}{N_T}\right] p_{n-1}$$

$$n = N_E\text{: } \sum_{j=0}^{N_E} p_j = 1 \tag{7}$$

where $\gamma = \lambda/\mu$.

The B_{N_T} are then found from Eq. (5), and x_i from Eq. (4). Finally, Fig. 14 is prepared.

D. Elimination of Starting Values in Exponential Smoothing, *R. G. Chamberlain*

Exponential smoothing is a technique for estimating the parameters of a model fit to a non-stationary discrete time series.

In its simplest form, the model of the underlying process is that successive data points, denoted x_t, are equal to a (possibly slowly varying) constant C plus a stochastic variate ϵ_t with zero mean, constant variance, and no serial correlation. That is,

$$x_t = C + \epsilon_t$$

With exponential smoothing, the estimated value of the constant $\hat{C}$ is obtained by a smoothed series,

$$\hat{C} \equiv S_t = ax_t + (1-a)S_{t-1}, \qquad 0 < a < 1 \tag{1}$$

This procedure has several difficulties: (1) a poor value of S_0 can cause a severe starting transient; (2) there is no provision for a variable time between samples; and (3), in consequence, there is no explicit way to deal with missing data points.

Equation (1) can be solved for S_t to give

$$S_t = a\sum_{i=0}^{t-1}(1-a)^i x_{t-i} + (1-a)^t S_0$$

Thus, past data are discounted exponentially (or, if you prefer, geometrically, since i takes on only integer values). The initial value becomes negligible only after

$$(1-a)^t << 1$$

The above difficulties can be overcome by a slight generalization. Let x_k and S_k be, respectively, the observation and the smoothed series at time t_k. Further, let us discount past observations exponentially with their age (as before), but eliminate the starting value, so that

$$S_k = b_k \sum_{i=1}^{k} \alpha^{t_k - t_i} x_i, \qquad 0 < \alpha < 1$$

with b_k chosen so that a constant input ($x_i = C$ for all i) will produce the same constant as output ($S_k = C$). A little algebra will produce the following recursion relations:

$$B_1 = 0 \qquad S_1 = x_1$$

$$B_k = \alpha^{t_k - t_{k-1}}(1 + B_{k-1}) \qquad S_k = \frac{x_k + B_k S_{k-1}}{1 + B_k} \tag{2}$$

where B_k is not the same number as b_k.

Equations (2) would be used if data points are not uniformly spaced or if one or more data points were missing. If data points are uniformly spaced, $t_k - t_{k-1} = \Delta t$ for all $k > 1$. Then,

$$B_k = \frac{\alpha^{\Delta t} - \alpha^{k\Delta t}}{1 - \alpha^{\Delta t}}$$

When k is large enough so that $\alpha^{(k-1)\Delta t} << 1$, B_k is constant, and

$$S_k = ax_k + (1-a)S_{k-1}$$

as in Eq. (1). The smoothing constants a and α are related by $a = 1 - \alpha^{\Delta t}$.

The generalization suggested here has not been applied to more complex exponential smoothing models (Refs. 1 and 2) but there appears to be no reason to doubt its feasibility.

References

1. Brown, R. G., *Smoothing Forecasting and Prediction of Discrete Time Series*. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1963.
2. Pegels, C. C., "Exponential Forecasting: Some New Variations," *Management Science*, Vol. 15, No. 5, Jan. 1969.

NASA — JPL — Coml., L.A., Calif.